Key concepts, words and tips will work as linkers = 🔗

The syllabus is divided into six topic areas:

- Basic economic ideas and resource allocation
- The price system and the microeconomy
- Government microeconomic intervention
- The macroeconomy
- Government macroeconomic intervention
- International economic issues

The key concepts of the Cambridge 9708 syllabus:

The **key concepts** are designed to help you develop a deeper understanding of the subject. They can open up new ways of thinking about, understanding or interpreting economic terms & theories.

Key concept	Description
Scarcity and choice	In economics, there is a fundamental problem: resources are scarce but wants are unlimited. Therefore, there is always a choice required between competing possible uses for the resources. There is therefore an opportunity cost in making this choice, such as when a decision is made to use a particular piece of land for agricultural or industrial use.
The margin and decision making	In economic theory, decisions by consumers, firms or governments will be based on choices taken at the margin. For example, firms will produce up to the point where the revenue generated by an extra unit of output is equal to the cost of producing it. However, economic decision making can also be based on values or ethical judgements, such as in relation to the need to try to reduce the extent of pollution.
Equilibrium and disequilibrium	Individual markets, as well as the whole economy, move in and out of equilibrium, constantly altering the allocation of scarce resources. For example, an individual market will be in a state of equilibrium when there is no excess demand or excess supply in that market.
Time	Economic conditions can change over a period of time. Time periods can vary between the short run, the long run and the very long run. Individuals, firms, markets and governments can respond to these changes in economic conditions in different ways, depending on the particular time frame — for example, trading off a cost in the present for a benefit in the future, such as when a new road is built.
Efficiency and inefficiency	It is possible for individual markets and the economy as a whole to be both efficient and inefficient in different ways when using scarce resources. For example, allocative efficiency occurs where price is equal to marginal cost.
The role of government and issues of equality and equity	There will be a trade-off between, on the one hand, freedom for firms and individuals in unregulated markets and, on the other hand, greater social equality and equity through different forms of government regulation of individuals and markets — for example, when governments intervene to bring about a more equal distribution of income and wealth in an economy, such as through a progressive tax.
Progress and development	Economics is concerned with how societies can make progress in measurable money terms and also how they can develop in a more normative sense in relation to standards of living — for example, the increasing emphasis that is now placed on sustainable development in many economies in the world.

Key skills in economics:

Economics involves several **key skills**, including the following:

» **Application**: you need to be able to apply economic theories and concepts to real-world examples.

» **Analysis**: you need to be able not simply to describe something, but also to analyse it in order to demonstrate that you clearly understand its essential features.

» **Evaluation**: you need to be able to demonstrate the skill of evaluation and critical thinking in particular situations. You should be capable of weighing up different arguments, prioritising certain options or policies and making logical and reasoned judgements.

» **Numerical skills**: you need to be able to calculate percentages, percentage changes and averages, and interpret index numbers.

» **Diagrams**: you need to be able to draw diagrams accurately and label them correctly to support a description, explanation or analysis.

» **Links between different topics**: you need to understand the links between different topics.

» **Problem solving**: you need to be able to analyse the causes and consequences of a problem and to make sensible suggestions about possible solutions.

» **Interpretation of data**: you need to be able to analyse given data critically, such as being able to identify a particular trend or to understand the limitations of the data.

The command words used in the economics exam:

Directive or command	Meaning of the directive or command word
Analyse	Examine in detail to show meaning, identify elements and the relationship between them
Assess	Make an informed judgement
Calculate	Work out from given facts, figures or information
Comment	Give an informed opinion
Compare	Identify/comment on similarities and/or differences
Consider	Review and respond to given information
Define	Give precise meaning
Demonstrate	Show how or give an example
Describe	State the points of a topic/give characteristics and main features
Discuss	Write about issue(s) or topic(s) in depth in a structured way
Evaluate	Judge or calculate the quality, importance, amount, or value of something
Explain	Set out purposes or reasons/make the relationships between things evident/provide why and/or how and support with relevant evidence
Give	Produce an answer from a given source or recall/memory
Identify	Name/select/recognise
Justify	Support a case with evidence/argument
Outline	Set out main points
State	Express in clear terms

As level:

1 Basic economic ideas and resource allocation

1.1 Scarcity, choice and opportunity cost

The fundamental economic problem and scarcity

- **wants:** items that are not essential for survival (e.g., a new car or television)
- **needs:** items that are essential for survival (e.g., food or shelter)
- **resources:** the inputs available to an economy for use in the production of goods and services
- **economic problem:** a situation where there are not enough resources to satisfy all human needs and wants
- **opportunity cost:** the benefit forgone from not choosing the next best alternative

» Scarcity refers to the fact that at any moment in time, the output that an economy is able to produce will be limited by the resources and technology available. People's **wants** and **needs**, however, will always exceed the **resources** available to satisfy them — in other words, these wants and needs are unlimited. This is known as the fundamental **economic problem**.
» As a result of this condition of scarcity, choices must be made.
» In all economies, therefore, there is an inevitability of choice at all levels of decision making — at the level of the individual, the firm and the government.

This focus on choice stresses the need to recognise the implications not only of choosing one thing, but also of not choosing something else. **Opportunity cost** is the benefit forgone from not choosing the next best alternative. An example is using a piece of land for farming purposes rather than building a factory on it.

🔗 It is important that candidates fully understand the difference between a want and a need, and can clearly demonstrate this understanding to the examiner.

KEY SKILL:

Application: an example of a *want* would be a new car or television and an example of a *need* would be food or shelter.

The need to make choices at all levels

As a result of the condition of scarcity, choices must be made. This could be at the level of:

» Individuals
» Firms
» Governments

The nature and definition of opportunity cost, arising from choices

The focus on choice stresses the need to recognise the implications not only of choosing one thing, but also of *not* choosing something else. This is known as opportunity cost.

🔗 Candidates sometimes define opportunity cost as the benefit that is forgone (or sacrificed) as a result of taking a decision. But it is not the result of any random choice; it is the cost of the next best alternative forgone.

🔗 **Scarcity and choice:** the fundamental problem in economics is that resources are scarce and wants are unlimited, so it is always necessary to make a choice between competing uses for the resources and there is always an opportunity cost in making this choice.

The basic questions of resource allocation

The emphasis on choice focuses on three basic economic questions:
» What to produce
» How to produce
» For whom to produce

The three basic economic questions are solved in different ways in various economies — in other words, resource allocation can be approached through different systems or *mechanisms*.

🔗 Candidates should emphasise the importance of needing to make a choice as a result of the condition of scarcity. Although choice can apply to various areas of economic activity, these three economic questions are the most fundamental ones.

KEY SKILL:

Problem solving: Understanding the potential missed opportunities forgone by economic agents when choosing one policy over another allows for better decision making.

1.2 Economic methodology

Economics as a social science

A social science can be defined as the scientific study of human society. Can economics be regarded as a social science?
» Economics is social in the sense that it studies different aspects of human behaviour and, in particular, the choices that humans make.
» Economics is a science in the sense that it uses an organised system of theories and facts capable of making verifiable predictions.
» Economics can therefore be regarded as a social science because it uses scientific methods to establish theories that can help explain the behaviour of individuals, groups and organisations in societies.

Positive and normative statements

- **positive statement:** a statement that is factual and objective
- **normative statement:** a statement that is subjective and expresses a value judgement
- **value judgement:** an opinion that reflects a particular point of view

It is important in economics to be able to distinguish between two different types of statements — positive statements and normative statements:

» A **positive statement** is one that can be checked against the facts to decide whether it is true.
» A **normative statement**, on the other hand, reflects the norms or values of the person expressing the statement — such a statement will involve a **value judgement** and will reflect someone's personal opinions. Normative statements often include the words 'should' or 'ought to'. The distinction between facts and value judgements is therefore very important in economics.

🔗 Candidates should understand that economics is one of the social sciences, so positive statements play an important role in the subject, offering an objective approach, whereas, in contrast, normative statements are more subjective and are reflections of value judgements.

The meaning of the term 'ceteris paribus'

- **ceteris paribus:** a Latin term that literally means 'other things being equal'
- **economic law:** an economic theory put forward by economists (e.g. the laws of demand and supply)
- **microeconomics:** the study of the behaviour of relatively small economic units (e.g. particular individuals, households or firms)
- **macroeconomics:** the study of economics at the national and international levels

» Although economics is one of the social sciences, with many aspects of the subject involving scientific analysis, it is not really possible to study human behaviour under laboratory conditions.
» However, economic theory does assume that certain aspects of human behaviour can be held constant.
» This assumption of **ceteris paribus**, that other things are equal, means that economists can analyse one aspect of human behaviour at a time. For example, in this way it has been possible to put forward **economic laws** of demand and supply. These economic theories have been put forward in relation to both **microeconomics** and **macroeconomics**.

🔗 Candidates should appreciate that it is virtually impossible to keep all variables constant, and this is why economists use the concept of ceteris paribus to indicate the idea of 'everything else being held constant'. This idea can be brought into a number of answers, such as showing the relationship between changes in the price of a product and changes in the demand for that product. If ceteris paribus applies, all other possible influences, such as changes in income, can be assumed to be constant.

The importance of the time period

- **short run:** the time period when it is not possible to change all of the factors of production
- **long run:** the time period when it becomes possible to change all of the factors of production
- **very long run:** the time period when technical progress is no longer assumed to be constant, as is the case in the short run and the long run, and the conditions of supply in an industry can be affected, e.g. by the impact of a new invention

Economists, when analysing economic behaviour, distinguish between three different time periods:

» **Short run:** this refers to that time period in which only certain factors of production can change. These are known as 'variable factors'. In the short run it is not possible to change the 'fixed factors'. For example, in the short run it may be possible to change labour, but the same capital will need to be used.
» **Long run:** this refers to that time period when the inputs of all factors of production can be changed — for example, it will be possible to vary *both* labour and capital in the long run. It is not possible to define exactly how long the short run or the long run is because it will vary depending on the particular circumstances.
» **Very long run:** this refers to that time period when supply conditions can change because of technical progress. In both the short run and the long run, technical progress is assumed to be held constant. In the very long run, however, technical progress can change — for example, as a result of a new invention in a particular industry — and this will have an effect on the supply conditions in that industry.

🔗 **Time:** economic conditions change in different time periods, such as the short run, the long run and the very long run. Individuals, firms, markets and governments are able to respond to these changes in different ways depending on the time frame.

1.3 Factors of production

The nature and definition of factors of production

- **primary sector:** production that takes place in agriculture, fishing, forestry, mining, quarrying and oil extraction
- **secondary sector:** production that takes place in manufacturing, construction and energy
- **tertiary sector:** production that takes place through the provision of services
- **land:** the factor of production that includes all the gifts of nature, or natural resources, that can be used in the process of production (e.g. minerals, forests and the sea)
- **labour:** the factor of production that includes all the human effort that goes into the process of production, both mental and physical
- **capital:** the factor of production that includes all the human-made aids to production (e.g. tools, equipment and machinery)
- **enterprise:** the factor of production that refers to taking a risk in organising the other three factors of production
- **entrepreneur:** the individual who takes a risk in combining the factors of production

Production in an economy can take place in three sectors:

» **Primary sector:** this is the extractive sector, where minerals are taken from the ground, and is concerned with production in areas of an economy such as farming, fishing, forestry, mining and quarrying.
» **Secondary sector:** this is the manufacturing and construction sector, working with the resources that have been extracted in the primary sector, and is concerned with areas of an economy such as car production and the construction of airport runways.
» **Tertiary sector:** this is the services sector and is concerned with wide areas of economic activity such as banking, insurance, tourism, teaching, medicine and the law.

There are four factors of production:

» **Land:** this refers to all the natural resources that can be used in the process of production. It can include farmland, forests, lakes and rivers and all the mineral deposits of a country, such as coal or oil.
» **Labour:** this refers to all the human input into the process of production. It refers not just to the people themselves, but to their skills, training, education and qualifications. It can also be referred to as ‘human capital’ or ‘intellectual capital’.
» Physical **capital:** this refers to the human-made aids that can be used in the process of production. It can refer to equipment, machinery and factories.
» **Enterprise:** this refers to the factor that brings the other factors of production together to produce products. The individual who combines the other factors of production, and takes a risk in doing so, is an **entrepreneur**.

🔗 Candidates often confuse the use of the term ‘capital’ as a factor of production with another use of the term to refer to money. It is important that these two meanings of the term are carefully distinguished.

The difference between human capital and physical capital

- **human capital:** the skills, knowledge and experience possessed by a population in terms of their value or cost to a business or an economy
- **physical capital:** the tangible, human-made objects that a business uses to produce goods and services (e.g. tools, machinery and equipment)

» **Human capital** refers to the human component of production — that is, the talent, knowledge, abilities, training, education and skills of the labour force.
» **Physical capital** refers to the non-human resources used in the production of goods and services — for example, the tools, equipment, plant, buildings and machinery.

🔗 Candidates often confuse the terms 'human capital' and 'physical capital'. It is important that these two terms are clearly distinguished.

The rewards to factors of production

- **rent:** the price paid for the use of land
- **wage:** the reward to labour based on the number of hours worked multiplied by an hourly rate of pay
- **salary:** the reward to labour on an annual basis
- **interest:** the reward for parting with liquidity; the reward to capital for the use of the human-made aids to production
- **profit:** the reward to enterprise, defined as the difference between total revenue and total costs

The rewards to the factors of production are as follows:

» **Rent:** the reward to land.
» **Wages** or **salaries:** the reward to labour.
» **Interest:** the reward to capital.
» **Profit:** the reward to enterprise; many enterprises aim for *profit maximisation*.

Division of labour and specialisation

- **specialisation:** the process whereby individuals, firms and economies concentrate on producing those products in which they have an advantage
- **division of labour:** the way in which production is divided into a sequence of specific tasks which enables workers to specialise in a particular type of job
- **Adam Smith:** one of the founding fathers of economics (1723–90) and author of *The Wealth of Nations*, published in 1776

Specialisation refers to a process of concentration on a particular aspect of production:

» A car assembly line is a good example of the way in which a manufacturing process can be broken down into a sequence of specific tasks. Workers will concentrate on, or *specialise* in, these particular tasks, giving rise to a **division of labour**.
» One of the first studies of this process was by the Scottish economist **Adam Smith**, who described in his book *The Wealth of Nations* (1776) how division of labour in a pin factory enabled a great many more pins to be produced than if each worker tried to do everything him- or herself.

KEY SKILL

Application: Adam Smith pointed out that the process of producing pins involved 18 specific operations. If one person did all of these, that person would be able to produce 20 pins a day. However, if division of labour was applied, it would be possible for each worker to produce 4,800 pins a day.

The role of the entrepreneur in contemporary economies

- **enterprise culture:** an economy in which taking a risk in the production of new products is encouraged in the hope of making a profit

Entrepreneurs play a crucial role in contemporary economies, performing two key functions:

» **Organisation:** entrepreneurs are responsible for organising and coordinating the other factors of production — land, labour and capital — to produce goods and services.
» **Risk:** entrepreneurs take a risk in performing this organisation and coordination function; this arises from the uncertainty that will be a feature of any initiative they take. Although there are many famous entrepreneurs in the world, who have had success in a number of different business ventures, there are many others who have failed.

Contemporary economies have provided many opportunities for the development of an **enterprise culture**. This is where people are imaginative and creative, and are willing to take risks in order to gain profit.

There are a number of ways in which a government could encourage the development of an enterprise culture, including:

» Supporting business start-up programmes
» Encouraging venture capital financing (this is where private investors provide finance to start-up businesses that are believed to have good long-term growth potential)
» Providing grants to support research and development
» Policies to promote competition in markets, such as deregulation
» Development of appropriate education and training to improve the quality of human capital
» Financial support for the development of technology parks and the fostering of innovation
» Favourable tax treatment for start-up businesses in the form of tax incentives
» Reducing administrative burdens, such as less ‘red tape’ in the form of excessive bureaucratic paperwork

1.4 Resource allocation in different economic systems

Decision making and resource allocation in market, planned and mixed economies

- **allocative mechanism:** a method of taking decisions about the different uses that can be made of factors of production

An **allocative mechanism** is needed for deciding how economic goods that are scarce are produced and consumed.

There are three different types of allocative mechanism:

1. **Market** economies 2. **Planned** economies 3. **Mixed** economies

🔗 Although an allocative mechanism is necessary to allocate economic goods, *free goods* that are in sufficient supply to satisfy demand do not need an allocative mechanism.

🔗 Candidates should understand that every country in the world (and there are over 200 countries) will allocate its scarce resources in different ways. This range of allocative mechanisms is so broad that economists have focused on three main types: market economies, planned economies and mixed economies.

1. **Market economies:**

- **market economy (or market system):** an economy where decisions about the allocation of resources are taken through the price mechanism
- **market:** a way in which buyers and sellers come together to exchange products

In a **market economy**, the allocation of resources is left to the **market** forces of demand and supply, operating through the price mechanism.

Advantages and disadvantages of the market economy

Advantages of the market economy	Disadvantages of the market economy
• Decisions are made by individual consumers, who act in their own self-interest, i.e. seek to maximise their utility or satisfaction when they consume a product. • Decisions are made by individual producers, who act in their own self-interest, i.e. seek to maximise their profits. • The use of the price mechanism to allocate resources (referred to as 'the invisible hand' by the Scottish economist Adam Smith) means that there is no need for any government intervention in the allocation of resources. • Competition between firms can lead to greater efficiency.	• Some products will be underprovided and under consumed in a market economy; these are known as merit goods (e.g. education and healthcare). • Some products will be overprovided and overconsumed in a market economy; these are known as demerit goods (e.g. alcohol and tobacco). • Some products will not be provided or consumed at all in a market economy because it would be impossible to charge a market price for them; these are known as public goods (e.g. defence and lighthouses). • Income and wealth disparities can be very significant.

2. Planned economies

- **planned (or command) economy:** an economy where decisions about the allocation of resources are taken by the state

Planned economies, aka **command economies**, involve the allocation of scarce resources through government intervention with no (or very little) scope for market forces to operate.

Advantages and disadvantages of the planned economy

Advantages of the planned economy	Disadvantages of the planned economy
• Government intervention in the allocation of resources means it can take decisions in the national interest (e.g. it can prevent the production of socially undesirable products such as drugs). • The government can intervene to bring	• A system with such a large amount of government influence and control will tend to be bureaucratic and, as a result, may be inefficient. • The lack of competition and the lack of the profit motive mean that products are

about a more equitable distribution of income and wealth.	often of poor quality with consumers having little choice.

3. Mixed economies

- **mixed economy:** an economy where the allocation of resources is decided both by market forces and by the state

A **mixed economy** combines elements of both market economies and planned economies — in other words, there is some degree of state ownership and state intervention, but in many areas of the economy market forces will be allowed to operate.

It could be argued that all economies today are, to some extent, mixed economies. However, there are large differences between, say, China, where the government still plays an important role in the allocation of resources, and the USA, where the government has only a limited role in the allocation of resources.

KEY SKILL

Evaluation: you need to be able to evaluate the strengths and weaknesses of the different types of allocative mechanism, coming to a judgement as to which is preferable and why. For example, a strength of a market economy is that there is no, or very little, government intervention, but a weakness is that without government intervention, there are likely to be many examples of market failure. Therefore, if a market is uncompetitive, or there is a high level of market failure, government intervention may be necessary to increase the degree of competition and reduce the level of market failure in the economy.

🔗 Candidates need to demonstrate they understand that the degree of mixture in any economy is not static. For example, since the credit crunch began in 2007, a number of banks in many countries have either been brought under complete state ownership or been given financial assistance by government to remain in business. One bank in the UK, NatWest, became 84% state owned in 2008 and this bank was still 51% state owned in 2022.

Transitional economic systems

- **transitional economy:** an economy that was previously a command or planned economy and which is now allowing a greater degree of scope for market forces to operate

A number of economies are going through a period of change where the extent of central planning is being reduced and market forces are being allowed to have a greater degree of influence. China and Cuba are examples of such a **transitional economy**. There are, however, possible problems associated with transition.

Problems of transitional economies

Unemployment	A planned economy is generally better able to keep down the rate of unemployment in an economy; when there is a move towards greater reliance on market forces, the rate of unemployment in an economy is likely to increase because, in a market economy, firms aim to maximise profits and this may lead them to reduce costs of production, possibly by laying off some workers.

Inflation	In a planned economy, the state controls prices so it is easier to keep down the rate of inflation; when prices are determined by the free-market forces of demand and supply, it is more difficult to control prices and so inflation is more likely.
Output	In a planned economy, it is possible for the state to support inefficient firms and industries; when state support is ended, such firms and industries may not be able to compete and so output could fall.
Welfare	A planned economy is able to provide housing and healthcare to everyone; with the introduction of market forces, there may be a fall in welfare provision and this may have a detrimental effect on levels of productivity in the economy.

🔗 Candidates should recognise that transitional economies can vary a great deal, depending on the degree of change or transition that has taken place. Some of these economies will still be similar to a planned economy, with only a small degree of private sector involvement. On the other hand, other economies will have moved away from a planned economy towards more of a market economy. It should also be understood that such economies are changing rapidly, and a great deal of change can have taken place in a short period of time.

1.5 Production possibility curves

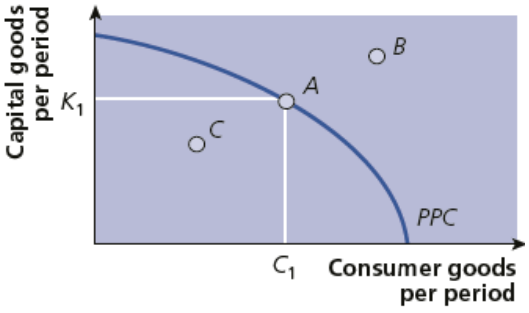
The nature and meaning of a production possibility curve (*PPC*)

- **production possibility curve (or frontier):** a graphic representation showing the maximum combination of goods or services which can be produced from given resources and with a constant state of technology

A **production possibility curve** (or **production possibility frontier**, as it is sometimes called) shows the different combinations of products that can be produced if an economy is working at full capacity.

The shape of the curve: constant and increasing opportunity costs

The shape of the curve shows that there are a number of different combinations of products that can be produced. It is drawn as a curve rather than as a straight line because not all factors of production are equally efficient.



F1.1 A production possibility curve

» The production possibility curve (*PPC*) in Figure 1.1 shows the combination of capital goods (shown on the vertical axis) and consumer goods (shown on the horizontal axis) that an economy can produce in a particular period of time with the existing economic resources available.
» Point *A* shows one possible combination of outputs, where the economy produces *K*1 capital goods and *C*1 consumer goods.
» Any movement along the curve from point *A* shows that the production of more of one type of good leads to the production of less of the other (thus illustrating the concept of *opportunity cost*).
» Point *C*, which is inside the *PPC*, shows that the economy is not using its resources efficiently and there is some unemployment of resources. Output of both capital and consumer goods is lower than it could be.

🔗 **Scarcity and choice:** the fundamental problem in economics is that resources are scarce and wants are unlimited, so a choice is always required between competing uses for the resources and an opportunity cost in making this choice.

Constant and increasing opportunity costs

» It has already been stated that a production possibility frontier is drawn as a curve, rather than as a straight line, because not all factors of production are equally efficient.
» It is therefore necessary to distinguish between constant and increasing opportunity costs. If it were possible to move from one point on the production possibility curve to another, with an equal sacrifice of resources, then this would indicate a situation of constant opportunity costs.
» However, there will come a time when this is not the case. Increasing opportunity costs mean that an ever-increasing amount of one product will need to be sacrificed to produce more of the other product.
» The reason is that different factors of production have different qualities. As a result of this, the production possibility frontier changes shape slightly as it approaches each axis.
» For example, in Figure 1.1, this is most clearly seen as the *PPC* gets closer to the horizontal axis, showing the consumer goods produced per period of time.

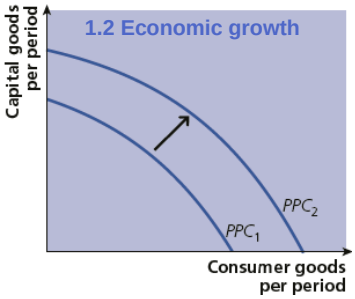
KEY SKILL

Analysis: a production possibility frontier is drawn as a curve because of the existence of the law of diminishing marginal returns (see Chapter 7, section 7.5). This states that employing an additional factor of production will eventually cause a relatively smaller increase in output.

🔗 **The margin and decision making:** the shape of the production possibility frontier as a curve illustrates the importance of decisions taken at the margin, given that resources are not equal substitutes for each other.

The causes and consequences of shifts in a *PPC*

- **economic growth:** an increase in the national output of an economy over a period of time, usually measured through changes in gross domestic product



» Point *B* in Figure 1.1, which is outside the *PPC*, is unreachable at the present time given the resources that the economy currently has.
» However, over a period of time it is possible for there to be **economic growth** resulting from the availability of more resources and/or the more productive use of resources, and this would enable point *B* to be reached

Economic growth enables an economy to produce more of both capital and consumer goods.
It refers to a situation where there is an expansion in the productive capacity or potential output of an economy.
This is shown in Figure 1.2 by a rightward shift of the *PPC* from *PPC*1 to *PPC*2. Of course, if there were a decrease in the quantity and/or quality of resources in an economy, this would lead to a leftward shift of a *PPC* from *PPC*2 to *PPC*1.

🔗 It is important that candidates understand the difference between a *movement along*, and *a shift* of, a production possibility curve:
» A movement along a curve indicates the different combinations of two goods that could be produced from the given resources in an economy.
» A shift of a curve to the right would indicate an expansion in the productive potential or capacity of an economy, allowing more of both goods to be produced.

KEY SKILLS

Diagrams: it is important that candidates understand the correct labelling of the two axes of a production possibility curve, especially when compared with the labelling of demand and supply diagrams in Chapter 2. The two axes of a *PPC* are labelled as particular goods or types of goods (e.g. capital goods and consumer goods in Figures 1.1 and 1.2). This is different from labelling the two axes *P* and *Q* in demand and supply diagrams.

Diagrams: it is important that a *PPC* is drawn so that it touches both axes. This is because it is assumed that all of an economy's resources will be used to produce one or other of the two products shown in the *PPC* diagram.

The significance of a position within a *PPC*

» It is important to understand the significance of a position within a *PPC*.
» point *A* is where an economy is using its resources efficiently, but point *C*, within a *PPC*, is where an economy is using its resources inefficiently.
» At this point, not all resources are being utilised and the output of both products is lower than it could be if all resources were being used.

1.6 Classification of goods and services

The nature and definition of free goods and private goods (economic goods)

1. Free goods

- **free good:** a good that is not scarce and so does not require a market price to be attached to it

A **free good** is one which is consumed by people without a situation of scarcity arising — in other words, there is enough of the good to satisfy everybody. As the good is not scarce, it does not require a market. The supply of the good equals the demand for it at zero price. It takes no factors of production to produce a free good and so, there is no opportunity cost involved.

2. Private goods

- **private good:** a good that is bought and consumed by individuals for their own benefit
- **excludability:** a feature of private goods whereby people can be excluded from consuming a good
- **rivalry:** a feature of private goods whereby when a product is consumed by one person, it cannot be consumed by another

A **private good** (or economic good) is one which is consumed by an individual for their own private benefit. This applies to most products in an economy. Private goods have two important characteristics:

» **Excludability**: one key feature of a private good is that people can be excluded from consuming it.
» **Rivalry**: another key feature of a private good is that the consumption of it by one person reduces its availability for other people; there is rivalry in such a situation because consumers are in competition with other consumers to consume a particular product.

🔗 It is important that candidates can clearly distinguish between *private goods* and *public goods* in their examination answers on this topic. The key characteristics of a private good are rivalry and excludability.

KEY SKILLS

Analysis: the situation of scarcity does not apply in the case of a free good because there is enough of a good to satisfy everybody.

Application: examples of free goods include air and sunshine.

Application: examples of private or economic goods include food, clothing, cars and smartphones.

The nature and definition of public goods

- **public good:** a good that is non-rival, non-excludable and non-rejectable
- **non-excludability:** where the consumption of a product by one person does not exclude others from consuming the same product
- **non-rivalry:** where the consumption of a product does not prevent its consumption by someone else
- **free rider:** the idea that it would be impossible to charge people for using a good or service because it would be impossible to prevent someone who had not paid from benefiting
- **government expenditure:** the total of all spending by a government
- **non-rejectability:** where individuals cannot actually avoid the consumption of a public good, even if they want to

In contrast to private goods, **public goods** are provided by society as a whole so that everyone can benefit from them. Public goods have two important characteristics:

» **Non-excludability**: once a public good has been provided for one person, it is not possible to stop other people from benefiting from such a good (i.e. no one is excluded).
» **Non-rivalry**: as more people consume the public good, the benefit to those already consuming it is not reduced (i.e. consumption by one person does not prevent others from consuming it).

These products need to be provided by the state or the public sector because if they were provided by the private sector, it would be impossible to exclude someone who had not paid. This gives rise to the **free rider** problem.

For example, it would not be possible to provide street lighting through the private sector because it would be impossible to prevent someone who had not paid from benefiting from the service. When such products are provided by the public sector, they are part of **government expenditure** and are financed out of taxation.

In addition to being non-rival and non-excludable, public goods are also **non-rejectable**. This means that, even if a person does not want to be protected by their country's defence and police system, they are not actually able to reject it.

Whereas key features of a private good are that it involves rivalry and excludability, candidates need to emphasise in their answers that key features of a public good are that it is both non-rival and non-excludable.

KEY SKILL

Application: examples of public goods include street lighting, defence and police.

A. The nature and definition of merit goods

- **merit good:** a product that is rivalrous and excludable but, if left to a free market, would be likely to be underproduced and underconsumed
- **information failure:** where people lack the full information that would allow them to make the best decisions about consumption
- **market imperfection:** a feature of a market which does not perform perfectly because of a failure to make an optimal use of resources, necessitating government intervention
- **market failure:** a market imperfection which gives rise to an allocation of scarce resources which is not as efficient as it might otherwise have been

» A **merit good** is a particular type of private good. Examples include education and healthcare.
» Like other private goods, merit goods are both rival and excludable, but what distinguishes a merit good is that there is **information failure** which means that the good is likely to be underprovided and underconsumed if provided through the private sector.
» For example, people don't fully appreciate the value of a good education or good health. This could be regarded as a **market imperfection**.
» Without government intervention, it is likely that there would be **market failure** because the allocation of resources would be sub-optimal.
» Governments therefore intervene by providing such goods through the public sector, alongside private sector provision, so that those who would not or could not afford to consume them in the private sector will do so in the public sector.

🔗 Candidates sometimes get confused and describe merit goods as examples of public goods. They are not examples of public goods, but of private goods. Like all private goods, they are rivalrous and excludable.

Candidates also sometimes confuse a merit good with a free good, especially given that some merit goods are free at the point of consumption, such as entry to a particular lesson. A free good, however, is something completely different: it is where there is so much of a product that demand can be satisfied without the need for an allocative mechanism, and supply will equal demand at zero price (e.g. air).

B. The nature and definition of demerit goods

- **demerit good:** a product that is rivalrous and excludable but, if left to a free market, would be likely to be overproduced and overconsumed

» **Demerit goods** are the opposite of merit goods. Whereas merit goods would be underprovided and underconsumed in a free market, demerit goods would be overproduced and overconsumed in a free market.
» A demerit good is socially undesirable in some way: for example, alcohol and tobacco.
» The overproduction and overconsumption of demerit gods is a result of imperfect information by consumers. For example, they may not realise that alcohol and tobacco are bad for their health.
» Without government intervention, it is likely that there would be market failure because the allocation of resources would be sub-optimal.

🔗 It is important that candidates indicate clearly how a demerit good is fundamentally different from a merit good. Whereas a merit good is likely to be underproduced and underconsumed, a demerit good is likely to be overproduced and overconsumed in a free market.

KEY SKILL

Application: examples of merit goods include education and healthcare. Examples of demerit goods include alcohol and tobacco. Examples of public goods include street lighting and defence.

2 The price system and the microeconomy

2.1 Demand and supply curves

Effective demand

- **effective demand:** demand for a product that is backed by the ability and willingness to pay for it

Effective demand refers to that demand which can be supported by having the means to pay. In this situation, consumers must not just want a particular product, but also be willing and able to pay for it.

KEY SKILL

Analysis: the characteristics of effective demand need to be made very clear when the theory of demand is being explained (i.e. the fact that people must be willing and able to buy something at a particular price).

🔗 It is important that candidates demonstrate in their answers an understanding that demand needs to be *effective demand*. It is not enough that consumers want something; they have to be in a position to pay for it.

Individual and market demand and supply

1. Individual and market demand

A. Individual and market demand curves

- **demand:** the quantity of a product that consumers are willing and able to buy at a given price in a given period of time
- **law of demand:** a law (or theory) which states that there is an inverse relationship between the quantity demanded of a product and the price of the product, ceteris paribus

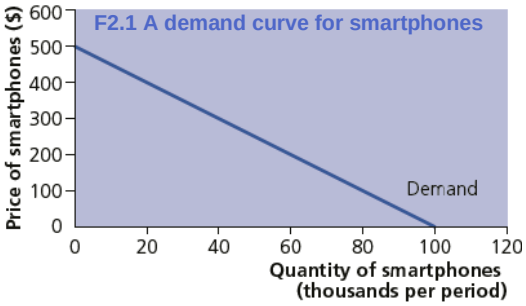
» **Demand** is the quantity of a product that consumers are willing and able to buy at a given price in a given time period.
» An individual demand curve shows the quantity of a product that a particular consumer is willing and able to buy at each and every price, ceteris paribus (i.e. with all other things unchanged).
» The individual demand curve will slope downwards from left to right, indicating that a consumer is more likely to buy a product at a lower price than at a higher price. This is known as the **law of demand**.

B. Aggregation of individual demand curves to give market demand

- **demand curve:** a curve that shows how much of a good or service will be demanded by consumers at a given price in a given period of time
- **demand schedule:** a table giving the quantities sold of a product at different prices, enabling a demand curve to be drawn from this information
- **derived demand:** where demand for the components of a product or for workers arises from demand for the final product

» A **demand curve** can be drawn for every consumer in a society for every product, but in economics it is more usual to focus on market demand curves.

» Market demand for a product is derived from bringing together (or aggregating) all the potential buyers of a product. It is the total quantity of a product that all potential buyers would choose to buy at a given price in a given period of time.
» A **demand schedule** can be produced for a particular product, such as smartphones.
» This schedule can then be plotted to give a market demand curve, as shown in Figure 2.1. The price of smartphones is shown on the vertical axis and the quantity of smartphones bought is shown on the horizontal axis.



The demand curve shows the relationship between price and the quantity demanded. It is downward sloping, indicating an inverse relationship between the price of a product and the quantity demanded of a product: that is, as the price falls, the demand rises.

Derived demand is where the demand for a component depends upon the final demand for a product that uses that component. For example, the demand for rubber is derived from the demand for car tyres. Derived demand can also be used in relation to the demand for workers — for example, the demand for bus drivers derives from people's demand for bus transport.

2. Individual and market supply

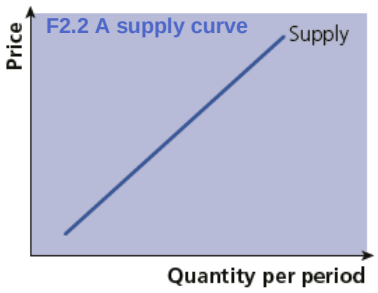
A. Individual and market supply curves

- **supply:** the quantity of a product that producers are willing to sell at a given price in a given period of time
- **law of supply:** a law (or theory) which states that there is a direct relationship between the quantity supplied of a product and the price of the product, ceteris paribus

Supply is the quantity of a particular product that firms are willing and able to sell at each and every price in a given time period, ceteris paribus (all other things unchanged). A firm's supply curve will slope upwards from left to right, indicating that a producer will be more likely to sell a product at a higher price than at a lower price. This is known as the **law of supply.**

B. Aggregation of individual firms' supply curves

- **supply curve**: a curve that shows how much of a good or service will be supplied by producers at a given price in a given period of time
- **supply schedule**: a table giving the quantities sold of a product at different prices, enabling a supply curve to be drawn from this information



supply curve can be drawn for every producer in an economy for every product, but in economics it is more usual to focus on market supply curves. Market supply of a product is derived from bringing together (or aggregating) all the potential suppliers of a product. It is the total quantity of a product that all potential sellers would choose to sell at a given price in a given period of time.

A **supply schedule** can be produced for a particular product, such as smartphones. This schedule can then be plotted to give a market supply curve, as shown in Figure 2.2.

The price of smartphones is shown on the vertical axis and the quantity of smartphones sold is shown on the horizontal axis. The supply curve shows the relationship between price and the

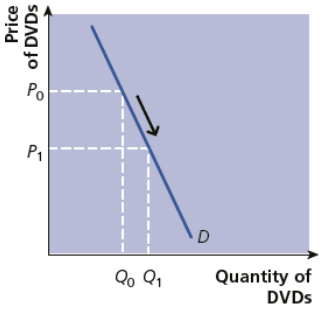
quantity supplied. It is upward sloping, indicating a direct relationship between the price of a product and the quantity supplied of a product: that is, as the price rises, the supply rises.

The determinants of demand

Price

- **change in quantity demanded:** where demand for a product changes as a result of a change in the price of the product; change in quantity demanded is shown by a movement along a demand curve
- **extension in demand:** when the quantity demanded of a product increases as a result of a fall in the price of the product, shown by a movement down the demand curve
- **contraction in demand:** when the quantity demanded of a product decreases as a result of a rise in the price of the product, shown by a movement up the demand curve

A major influence on the demand for a product is its price. Figure 2.3 shows that there is an inverse relationship between a change in the price of a product and the quantity demanded of a product, all other things unchanged (ceteris paribus).



F2.3 A movement along the demand curve

When it is only the price of a product that changes, the resulting **change in quantity demanded** can be shown on a demand curve by a movement *along* the curve.

» When the price of a product is reduced, for example, from *P*0 to *P*1, the quantity demanded goes up from *Q*0 to *Q*1.
This is represented by a downward movement along the demand curve, indicated in the diagram by the downwards arrow.
This is known as an **extension in demand**.
» If, on the other hand, the price of a product is increased, the quantity demanded falls and this would be shown as an upward movement along the demand curve.
This is known as a **contraction in demand**.

The determinants of supply

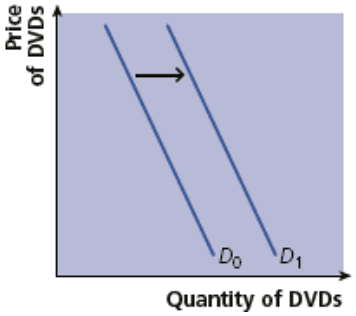
change in quantity supplied: where the supply of a product changes as a result of a change in the price of the product; change in quantity supplied is shown by a movement along a supply curve
extension in supply: when the quantity supplied of a product increases as a result of a rise in the price of the product, shown by a movement up the supply curve
contraction in supply: when the quantity supplied of a product decreases as a result of a fall in the price of the product, shown by a movement down the supply curve

Movements along a supply curve are determined by changes in the price of a product. Figure 2.4 shows the direct relationship between the price of a product and the quantity supplied of that product. This is why the supply curve is upward sloping. A movement up a supply curve is known as an **extension in supply** and a movement down a supply curve is known as a **contraction in supply**.

Causes of a shift in the demand curve

- **change in demand:** where there is a change in the conditions of demand, i.e. something other than a change in the price of a product; this is shown by a shift of a demand curve
- **composite demand:** the demand for a product that can be used for more than one purpose

Price is not the only factor that influences demand. If the ceteris paribus assumption is removed, it is possible to consider all the other factors that were previously being held constant. These other factors could include:



F2.4 A shift in the demand curve

» a change in the incomes of consumers
» a change in the price of a substitute product (a substitute product is one that could be used for the same purpose by consumers)
» a change in the price of a complementary product (a complementary product is one that is directly related to, and used with, another product)
» an advertising campaign
» a change in population
» a change in the tastes and preferences of consumers
» a lowering of interest rates, making borrowing more affordable
» a change in the weather, possibly associated with different seasons

When one of these other factors affects demand, the result is described as a **change in demand** and is shown by a *shift* of the demand curve.

In this diagram, there might have been an increase in incomes and/or an effective advertising campaign. The demand curve shifts to the right, from *D*0 to *D*1, as shown by the rightward arrow.

Composite demand refers to the demand for a product that can be used for more than one purpose. Stone, for example, could be used for building purposes and could also be used in the construction of roads; a particular piece of land could be demanded to build both shops and houses.

🔗 Candidates sometimes confuse movements *along* a demand curve and a *shift* of a demand curve. It is important that you understand what will cause a movement along a demand curve and what will cause a shift of a demand curve. A movement along a demand curve can only be caused by a change in the price of a product, whereas a shift of a demand curve can be caused by anything other than a change in the price of a product.

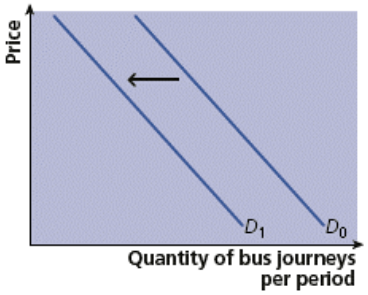
KEY SKILL

Diagrams: it is important to show clearly the direction of a shift in a demand curve: for example, by labelling the two demand curves *D*0 and *D*1 and by including an arrow to show the direction of the shift.

Normal and inferior goods

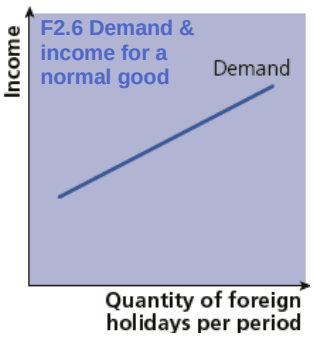
- **normal good:** a good for which the demand rises with an increase in income
- **inferior good:** a good for which the demand falls with an increase in income

» Figure 2.4 showed what usually happens when there is an increase in the incomes of consumers — more of the product is bought at every price and there is a rightward shift of the demand curve, showing an increase in demand. Such goods are called **normal goods**.
» However, it is possible that the demand for some goods and services *decreases* when there is an increase in incomes. For example, while there might be an increase in the demand for cars as a result of an increase in incomes, there might be a decrease in the demand for public transport, such as bus journeys.



F2.5 A shift in the demand curve following an increase in consumer incomes (an inferior good)

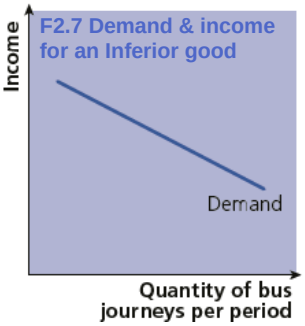
» This can be seen in Figure 2.5 where there is a leftward shift in the demand curve for bus journeys, showing a decrease in demand. Such goods are called **inferior goods**.



F2.6 Demand & income for a normal good

It is important to recognise that the demand for normal and inferior goods shows the relationship between a change in the quantity demanded and a change in income, not price. Figure 2.6 shows this relationship for a normal good.

🔗 Candidates need to ensure that they understand the difference between a *normal good* and an *inferior good* and can demonstrate this in their examination answers. A normal good is one where demand will increase as a result of a rise in income. An inferior good is the opposite: it is a good where demand will decrease as a result of a rise in income.



F2.7 Demand & income for an Inferior good

🔗 Candidates can sometimes confuse the effect of a change in *price* and a change in *income* in examinations. These two effects need to be clearly distinguished. For example, changes in the quantity demanded of normal and inferior goods take place in response to a change in a person's income, not to changes in the prices of the goods.

KEY SKILL

Evaluation: it needs to be recognised that what is described as a normal good and what is described as an inferior good may vary between different countries, or within one country at different historical time periods.

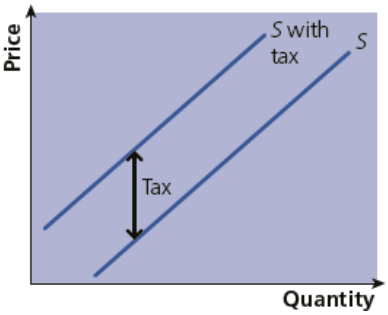
Causes of a shift in the supply curve

Price is not the only factor that influences supply. If the ceteris paribus assumption is removed, it is possible to consider all the other factors that were previously being held constant. These other factors could include the following :

Indirect taxes

- **indirect tax:** a tax that is imposed on expenditure; it is indirect in that the tax is only paid when the product on which the tax is levied is purchased

» A government may decide to impose an **indirect tax**, such as a sales tax, on a particular good or service. Examples are value added tax (VAT) and goods and services tax (GST).
» The effect of the imposition of such a tax can be seen in F2.8. In this case, the tax is a specific tax with a fixed amount of tax per unit, so the supply curve shifts upwards, parallel to the original supply curve.
» Figure 2.8 shows how the imposition of an indirect tax will affect the price of a product. As there is an upward movement of the supply curve, this will lead to an increase in price. In order to determine the exact price charged, it would be necessary to include a demand curve in the diagram.



F2.8 The effect of a sales tax on supply

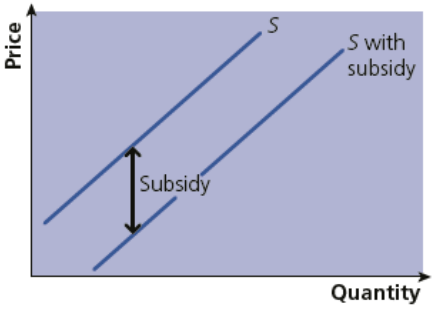
KEY SKILL

Analysis: if the indirect tax is an ad valorem tax, which adds a certain percentage on to the price, the supply curve will not shift upwards parallel to the original supply curve but will move further away from it.

Subsidies

- **subsidy:** an amount of money paid by a government to a producer, so that the price charged to the customer will be lower than would have been the case without the subsidy

» Whereas the imposition of a tax shifted the supply curve upwards to the left, a subsidy has the opposite effect. If a government pays firms a subsidy to produce a particular product, this will have the effect of reducing their costs and encourage firms to supply more output at any given price. This can be seen with the supply curve shifting downwards to the right.
» The effect of the subsidy, in shifting the supply curve to the right, will be a lowering of price. The actual price will be determined where the '*S* with subsidy' line intersects with the demand curve. The effect of the subsidy is that both producers and consumers may benefit.



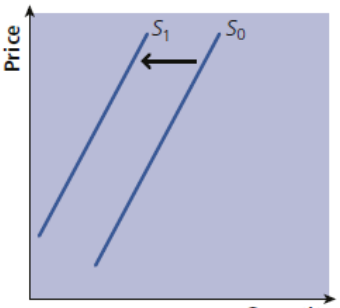
F2.9 The effect of a subsidy on supply

🔗 The distinction between the effect of a tax and the effect of a subsidy is another area that candidates often confuse in examinations. You need to remember that the effect of a tax is to shift the supply curve to the left, whereas the effect of a subsidy is to shift the supply curve to the right.

Production costs

An important influence on supply is the costs of production. If the costs of the inputs in the production process — that is, the costs of the factors of production — increase, then firms will be inclined to supply less output at any given price.

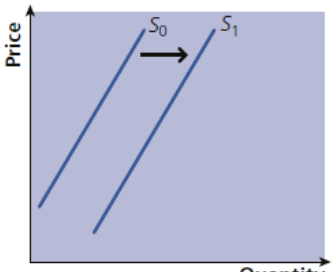
The increase in production costs causes the supply curve to shift to the left from *S*0 to *S*1.
The increase in costs can be seen by the vertical distance between *S*0 and *S*1.



F2.10 The supply curve Shifts to the left if production costs increase

Technology of production

Another important influence on supply is the technology of production. If the technology of production is improved, this means that firms will be able to produce more effectively than before.
improved technology leads to firms supplying a greater output at any given price. The supply curve shifts to the right from *S*0 to *S*1.



2.11 The supply curve shifts to the right if production costs fall

The prices of other goods

» There may be a degree of substitution on the supply side if the prices of different products change.
» In many cases, the factors of production that a firm has can have alternative uses, so a firm may be influenced by changes in the prices of different products to produce more of one product and less of another.
» A rise in the price of a product could increase its profitability, so a firm may decide to switch production towards this product. For example, car production companies are increasingly switching production towards electric cars.

Expected prices

A final influence on market supply relates to the expectations of firms about possible future prices. This is especially the case in those situations where the production process takes quite a long time. Firms will thus need to take supply decisions on the basis of expected prices in the future. This is often the case in agriculture.

The distinction between the shift in the demand or supply curve and the movement along these curves

» It has already been pointed out that it is important to distinguish between a movement along a demand curve and a shift of a demand curve.

» If there is a change in the market price of a product, and nothing else changes (i.e. assuming ceteris paribus), this will involve a movement along a demand curve. This shows how consumers react to a change in the price of a product. This can be seen in Figure 2.3.

» If the situation of ceteris paribus cannot be assumed, however, and there is a change in any of the other influences on demand, then the demand curve will shift. This could involve a shift to the left, showing a decrease in demand (e.g. as a result of a lowering of incomes) or a shift to the right, showing an increase in demand (e.g. as a result of an effective advertising campaign). The latter case was shown in Figure 2.4.

» It is also important to distinguish between a movement along a supply curve and a shift of a supply curve. If there is a change in the market price of a product, and nothing else changes (again assuming ceteris paribus), this involves a movement along a supply curve. This shows how firms react to a change in the price of a product and was shown in Figure 2.2.

» If the situation of ceteris paribus cannot be assumed, however, and there is a change in any of the other possible influences on supply, then the supply curve will shift because this will affect the willingness of firms to supply at any given price.

» As has already been indicated, this could involve a shift to the left: for example, as a result of the imposition of an indirect tax on the consumption of a product (see Figure 2.8) or an increase in production costs (see Figure 2.10). Alternatively, it could involve a shift to the right: for example, as a result of the introduction of a subsidy (see Figure 2.9) or an improvement in technology (see Figure 2.11).

🔗 The possible confusion between a movement along a demand curve and a shift of a demand curve has already been pointed out. The possibility of confusion also applies to supply.
You need to be absolutely certain that you understand the difference between a movement along a supply curve and a shift of a supply curve before taking the examination.

2.2 Price elasticity, income elasticity and cross elasticity of demand

The definition of price elasticity of demand (*PED*), income elasticity of demand (*YED*) and cross elasticity of demand (*XED*)

The concept of elasticity of demand refers to the responsiveness of demand to a change in one of its determinants, such as the price of a product, income or the price of another product.
There are three elasticities of demand:

1. **Price elasticity of demand (*PED*):** this measures the responsiveness of the demand for a product to a change in its price.

2. **Income elasticity of demand (*YED*):** this measures the responsiveness of the demand for a product to a change in income.

3. **Cross elasticity of demand (or cross-price elasticity of demand) (*XED*):** this measures the responsiveness of demand for a product to a change in the price of another product.

The formulae and calculation of price elasticity, income elasticity and cross elasticity of demand

Elasticities of demand are calculated by dividing the percentage change in the quantity demanded of a product by the percentage change in the determinant causing the change in demand.

The three elasticities of demand are as follows:

1. **Price elasticity of demand** measures the responsiveness of the demand for a product to a change in its price. It is calculated by the following formula:

$$\frac{\text{percentage change in the quantity demanded of a product}}{\text{percentage change in the price of a product}}$$

2. **Income elasticity of demand** measures the responsiveness of the demand for a product to a change in income. It is calculated by the following formula:

$$\frac{\text{percentage change in the quantity demanded of a product}}{\text{percentage change in income}}$$

3. **Cross elasticity of demand** or **cross-price elasticity of demand** measures the responsiveness of demand for one product to a change in the price of another product. It is calculated by the following formula:

$$\frac{\text{percentage change in the quantity demanded of good A}}{\text{percentage change in the price of good B}}$$

KEY SKILLS

Numerical skills: you need to be able to calculate *PED*. For example, if the price of a product increases by 20% and the quantity demanded decreases by 10%, *PED* = 10%/20% = 0.5. There should really be a minus sign before the 0.5 because it is a negative number: that is, there is an inverse relationship between the change in price and the change in demand. However, the minus sign is usually left out. This is because it is expected that there will be a negative or inverse relationship between quantity demanded and price.

Numerical skills: you need to be able to calculate a percentage change. For example, to convert 35/86 into a percentage, the numerator (35) is divided by the denominator (86) and the answer multiplied by 100, so 35 divided by 86 = 0.406976744 or 0.407. This is then multiplied by 100 to arrive at the answer of 40.7%.

Numerical skills: you need to be able to calculate *YED*. For example, if income increases by 5% and the quantity demanded of a product increases by 20%, *YED* = 20%/5% = 4.

Numerical skills: you need to be able to calculate *XED*. For example, if the price of a product increases by 10% and the demand for another product increases by 15%, *XED* = 15%/10% = 1.5.

The significance of relative percentage changes, the size and sign of the coefficient of the three elasticities

Price elasticity of demand

- **price elasticity of demand:** measures the degree to which a change in the price of a product leads to a change in the quantity demanded of the product

If the price of a good rises by 20%, and the quantity falls by 40%, then the **price elasticity of demand** is 40% divided by 20% = 2. There should really be a minus sign before the 2 because it is a negative number: that is, there is an inverse relationship between the change in price and the change in demand.

Income elasticity of demand

- **income elasticity of demand:** measures the degree to which a change in incomes leads to a change in the quantity demanded of a product

The **income elasticity of demand** for most products will be positive: that is, as incomes rise, the demand for products will rise. As we have seen, these are known as *normal goods*. However, the income elasticity of demand for some products will be negative: as incomes rise, the demand for products will fall. These are known as *inferior goods*.

KEY SKILL

Application: a necessity is a normal good with a positive *YED*, but the income elasticity of demand for such a product (e.g. an essential item of clothing) will be between 0 and 1. A luxury good, such as an expensive holiday, will also have a positive *YED*, but with an income elasticity of demand of more than 1.

Cross elasticity of demand

- **cross elasticity of demand (or cross-price elasticity of demand):** measures the degree to which a change in the price of one product leads to a change in the quantity demanded of another product
- **substitute goods:** goods which are possible alternatives (e.g. gas or electricity as a source of energy in a home); these goods have a positive cross elasticity of demand (i.e. a rise in the price of one of them will lead to an increase in the demand for the other)
- **complementary goods:** goods which are consumed together (e.g. printers and ink cartridges); these goods have a negative cross elasticity of demand (i.e. a rise in the price of one of them will lead to a decrease in the demand for the other)

The **cross elasticity of demand** will be positive if two goods are **substitutes**, such as tea and coffee. If good B increases in price, a number of people will switch to the substitute, good A, and so the demand for good A increases.

If the two goods are **complements**, such as printers and ink cartridges, the cross elasticity of demand will be negative. As the price of good B rises, fewer people will buy it and so fewer people will buy good A as well.

🔗 A number of candidates write the formulas for the three elasticities of demand the wrong way round in examinations. To avoid making this mistake, remember that in all three calculations (i.e. price, income and cross elasticity of demand) the percentage change in the quantity demanded is always on the top.

Descriptions of elasticity values

- **perfectly inelastic:** where a change in an independent variable has no effect on the quantity demanded (or supplied); the calculation will be zero and it is shown as a vertical straight line

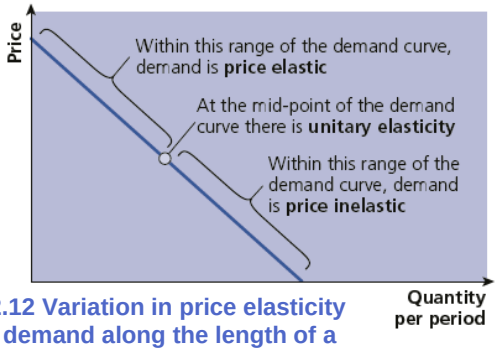
- **inelastic:** where the response of demand (or supply) is proportionately less than the change in the independent variable; the calculation is less than 1
- **unitary elasticity:** where the proportionate change in demand (or supply) is exactly equal to the change in the independent variable; the calculation will be equal to 1, it will be represented by a rectangular hyperbola (in the case of demand) and a movement up or down a demand curve will leave **total revenue** unchanged
- **elastic:** where the response of demand (or supply) is proportionately greater than the change in the independent variable; the calculation is greater than 1
- **perfectly elastic:** where all that is produced is bought/sold; the calculation is infinity and it is shown as a horizontal straight line
- **total revenue:** the total amount of income received from sales of a product, calculated as the number of units sold multiplied by the price of each unit

Elasticity of demand can vary from perfectly inelastic to perfectly elastic:

Elasticity	Figure
Perfectly inelastic	Zero
Inelastic	Greater than zero but less than 1
Unit elastic	1
Elastic	Greater than 1 but less than infinity
Perfectly elastic	Infinity

The variation in price elasticity of demand along the length of a straight-line demand curve

A straight-line demand curve does not indicate constant elasticity of demand along the entire length, except in the case of perfectly elastic and perfectly inelastic curves. The price elasticity of demand will, in fact, vary along the line.


Price
Within this range of the demand curve, demand is price elastic
At the mid-point of the demand curve there is unitary elasticity
Within this range of the demand curve, demand is price inelastic
Quantity per period

F2.12 Variation in price elasticity of demand along the length of a straight-line demand curve

The factors affecting elasticity of demand

The various factors affecting elasticity of demand can be seen in relation to the three different types of elasticity.

The factors affecting price elasticity of demand

There are a number of factors affecting price elasticity of demand, including the following:

» **Availability of substitutes:** the more substitutes that are available for a particular product, such as different types of tea or coffee, the more price elastic will be the demand.

» **Definition of the product:** the narrower the definition of a product, the more price inelastic will be the demand: for example, the demand for tea or coffee will be more inelastic than the demand for non-alcoholic beverages.

» **Amount spent on the product:** if the amount spent on a product is a relatively small percentage of a person's income, the demand is likely to be inelastic: for example, the amount spent on boxes of matches or on newspapers is likely to be a small percentage of weekly expenditure, so the demand for such products is likely to be relatively inelastic.

» **Time:** demand for a product is likely to be more inelastic in the short run than in the long run: for example, a decision to buy a new printer to replace one that has broken.

🔗 **Time:** the period of time can affect the price elasticity of demand for a product; it is likely to be more elastic in the long run than in the short run because it gives consumers more time to think about possible alternatives to a product.

🔗 A common error in examinations is to describe a particular good as elastic or inelastic. It is important that you avoid this mistake. It is not a good that is elastic or inelastic, but the *demand* for a particular good that is elastic or inelastic.

The factors affecting income elasticity of demand

There are a number of factors affecting income elasticity of demand, including the following:

» **Proportion of income that is spent on a particular good:** the demand for some products, such as matches, will not be very sensitive to a change in income because they are not very expensive; in these cases, income elasticity of demand will be virtually zero.

» **Definition of the product:** the income elasticity of demand for cars will be positive, but it may be negative for particular, cheaper, models of cars.

» **Economic development of a particular economy:** in some economies, a motorcycle may be regarded as a normal good, so the income elasticity of demand will be positive, but as the economy develops and more people can afford cars, the demand for motorcycles may fall, a motorcycle may start to be regarded as an inferior good and so the income elasticity of demand for motorcycles will be negative.

The factors affecting cross elasticity of demand

There are a number of factors affecting cross elasticity of demand. These include:

» Whether the relationship is between substitutes or complements (this would determine the *XED* sign and not the value of the *XED* coefficient)

» Whether the substitutes are close or weak substitutes or whether the complements are close or weak complements; the stronger the relationship between two products, the higher is the coefficient of *XED* (examples of strong complements are smartphones and apps, whereas examples of weak complements are shoes and polish)

» Whether there is any relationship at all between two products; unrelated products have a zero cross elasticity of demand

The relationship between price elasticity of demand and total expenditure on a product

There is a close relationship between *PED* and total expenditure on a product:
» If price elasticity of demand is less than unitary (i.e. it is inelastic), a fall in the price of a product causes a fall in total expenditure on the product and a rise in the price of a product causes a rise in total expenditure on the product. When demand is inelastic, price and total expenditure move in the same direction.

» If price elasticity of demand is more than unitary (i.e. it is elastic), a fall in the price of a product causes a rise in total expenditure on the product and a rise in the price of a product causes a fall in total expenditure on the product. When demand is elastic, price and total expenditure therefore move in opposite directions.

» If price elasticity of demand is unitary (i.e. equal to 1), a rise or fall in the price of a product causes no change in total expenditure on the product.

» If price elasticity of demand is perfectly elastic, the slightest rise in price will lead to zero demand for a product, so total expenditure will be zero.

» If price elasticity of demand is perfectly inelastic, a change in price will cause no change in the quantity demanded of a product, so the change in total expenditure will depend on whether there has been a fall in price or a rise in price.

The implications for decision making of price, income and cross elasticity of demand

Price elasticity of demand

Price elasticity of demand is very important to an understanding of business decisions, especially because of the link with revenue:

» **Price elastic:** if the demand for a product is price elastic, a business should lower the price of the product because more products will be bought, and this will produce a higher total revenue.

» **Price inelastic:** if the demand for a product is price inelastic, a business should increase the price of the product because, even though fewer items will be bought, the increased revenue from each product sold will offset this and therefore, total revenue will increase.

the link between price changes and revenue changes in relation to different price elasticities of demand:

Price elasticity of demand	For a price increase, total revenue will	For a price decrease, total revenue will
Inelastic	…rise	…fall
Unitary elastic	…stay the same	…. stay the same
Elastic	…fall	…rise

Income elasticity of demand

Income elasticity of demand is also important to an understanding of business decisions. Changes in an economy, and particularly changes in the level of incomes, can influence what a business is going to produce or stock. This is indicated by the following two examples:

» **Rising incomes:** if an economy is growing and incomes are rising, a business might want to move from inferior goods towards producing normal goods; this will influence the planning of businesses in the future, such as in relation to employment requirements.

» **Falling incomes:** if an economy is experiencing a recession and incomes are falling, a business will want to be producing or stocking products with a relatively low income elasticity of demand; for example, people will still want to buy food in a recession, but they are much less likely to want to buy expensive cars.

Cross elasticity of demand

Similarly, cross elasticity of demand is also important to an understanding of business decisions. This is indicated by the following two examples:

» **Substitutes:** in the case of a substitute, a firm would be able to estimate the effect on the demand for its product of a change in the price charged by another firm in the market, such as in relation to a change in the price of tea and the demand for coffee.

» **Complements**: a firm would be able to estimate the effect on the demand for a product if there was a change in the price of a complement; for example, a fall in the price of DVD players would be likely to lead to an increase in the demand for DVD players and, therefore, an increase in the demand for DVDs.

2.3 Price elasticity of supply

The definition of price elasticity of supply

- **price elasticity of supply:** measures the degree to which a change in the price of a product leads to a change in the quantity supplied of the product

Price elasticity of supply measures the responsiveness of the supply of a product to a change in its price.

The formula for, and calculation of, price elasticity of supply

Price elasticity of supply is calculated by the following formula:

$$\frac{\text{percentage change in the quantity supplied of a product}}{\text{percentage change in the price of a product}}$$

The significance of relative percentage changes and the size and sign of the coefficient of price elasticity of supply

Price elasticity of supply can range from perfectly elastic to perfectly inelastic (from infinity to zero):

» If a firm will supply any quantity of a good at the going price, then supply is perfectly elastic or infinite; a perfectly elastic supply curve will be horizontal.

» If the percentage change in supply is greater than the percentage change in price, then supply is price elastic; an elastic supply curve will always intersect with the price axis.
» If the percentage change in supply is equal to the percentage change in price, then supply is unitary elastic; any supply curve drawn from the origin will have a unit price elasticity of supply.

» If the percentage change in supply is less than the percentage change in price, then supply is inelastic; an inelastic supply curve will always intersect with the quantity axis.

» If a firm will supply only a fixed quantity of a good at the going price, and cannot increase or decrease the amount available, then supply is perfectly inelastic or zero; a perfectly inelastic supply curve will be vertical.

The sign of the coefficient of price elasticity of supply is always positive.

🔗 As with price elasticity of demand, make sure you remember that the change in quantity goes on top of the formula and the change in price on the bottom.

The factors affecting price elasticity of supply

- **stocks:** goods that have been produced, but that are unsold and stored for sale in the future (e.g. a firm that sells car tyres usually has considerable stocks of tyres to fit a wide range of cars)
- **perishability:** the length of time in which a product is likely to decay or go bad — the shorter the time, the more perishable the product (e.g. cheese usually has a sell-by date and a date by which it should be consumed)

There are a number of factors affecting price elasticity of supply, including the following:

» **Number of producers:** the greater the number of suppliers, the more likely it is for the industry to increase output in response to a price increase, so supply is likely to be relatively elastic.

» **Amount of stocks:** some products will be easier to stock than others, and this will make the supply of them relatively more elastic; but some products will be **perishable** and so more difficult to stock for long periods, making their supply less elastic (perishable products include meat, seafood and dairy produce).

» **Time period**: supply is likely to be more elastic over a longer period of time (see Figure 2.13), as this gives firms more time to invest in more factors of production and also gives more time for new firms to join the industry.

» **Existence of spare capacity:** the greater the degree of capacity in the industry, the easier it will be for firms to increase output if the price of products increases, and this is likely to make supply more elastic.

» **Length of the production period:** supply is usually more elastic in manufacturing than in agriculture because manufacturing usually involves a shorter production period than agriculture.

» **Degree of factor mobility:** the easier it is for economic resources to be transferred into the industry, the more elastic the supply is likely to be.

🔗 **Time:** supply is likely to be more elastic over a longer period of time as it gives more time for new firms to join an industry.

The implications for speed and ease with which firms react to changed market conditions

The various factors covered in the previous section give an indication of the speed and ease with which firms in an industry can respond to changed market conditions.

Figure 2.13 shows the relationship between the price elasticity of supply and the time period:
» Supply curve *S*s shows supply in the short run, when it will usually be more difficult to alter supply at relatively short notice and supply tends to be relatively inelastic.

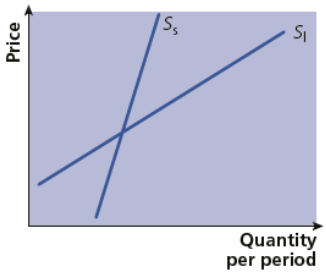
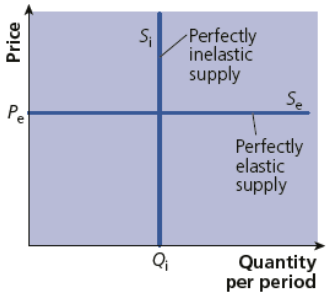


Figure 2.13 Short- and long-run supply

» Supply curve *S*l, however, shows supply in the long run, when firms are usually more able to increase production and so supply tends to be relatively elastic.

» It is possible that there may be a situation of perfectly inelastic supply at a particular moment in time; in this situation, it is not possible to increase supply, no matter how much price increases by. Examples are seats in a cinema or spaces in a parking lot.

» Agricultural products are a good example of perfectly inelastic supply at a particular moment in time, as it can take a number of years to bring such products to the market.



F2.14 Perfectly elastic and inelastic supply

» At the other extreme, it is possible that there may be a situation of perfectly elastic supply; in this situation, the firms in the industry would be willing to supply any amount of the product at a given price. For example, if resources are available, the supply of batteries by firms in the industry may become perfectly elastic.

KEY SKILL

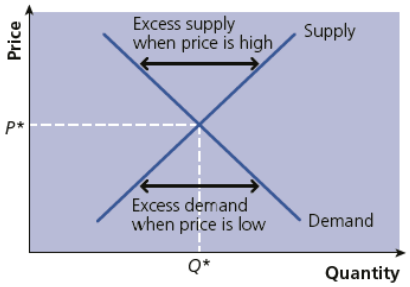
Analysis: you need to be able to analyse how price elasticity of supply can vary between firms in different industries, especially between those involved in the production of agricultural and manufactured products.

2.4 The interaction of demand and supply

The meaning of market equilibrium and disequilibrium

- **equilibrium:** a situation where the quantity demanded in the marketplace is exactly equal to the quantity supplied and there is neither excess demand nor excess supply in the market; sometimes referred to as a state of rest or balance or stability where there is no tendency to change
- **disequilibrium:** a situation where there is an imbalance between demand and supply in a market (i.e. there is either excess demand giving rise to a shortage or excess supply giving rise to a surplus)

Having considered both demand and supply, it is now necessary to bring them together to establish what is meant by 'market equilibrium'.



F2.15 Bringing demand and supply together

» Market equilibrium is shown in Figure 2.15.

» The downward-sloping demand curve and the upward-sloping supply curve cross at the equilibrium position of price *P** and quantity *Q**.

» If the price were higher than this, there would be excess supply and this would cause the price to move downwards to the equilibrium position.

» If the price were lower than this, there would be excess demand and this would cause the price to move upwards to the equilibrium position.

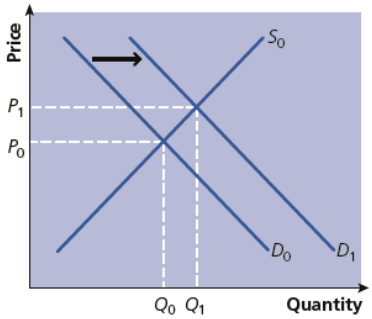
» If a situation of excess supply or excess demand were to exist for a period of time, this would be called **disequilibrium** until a position of equilibrium was eventually restored.

🔗 **Equilibrium and disequilibrium:** individual markets, and the economy as a whole, are always moving into and out of equilibrium, constantly altering the allocation of resources.

🔗 It is important that candidates can distinguish between a situation of *equilibrium* and one of *disequilibrium* in a market.

The effect of shifts in demand and supply on equilibrium price and quantity

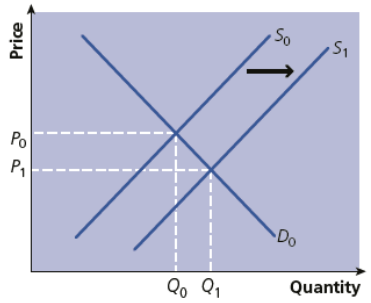
- **equilibrium price:** the price at which a market clears (this means that at this price, the quantity demanded equals the quantity supplied); the process of market clearing arises because the price is free to change and settle at the equilibrium level
- **equilibrium quantity:** the quantity at which a market clears, with consumers getting all they want at the equilibrium price and producers not being left with unsold products (i.e. there is no excess demand or supply)



F2.16 The effect of a shift of a demand curve to the right on equilibrium price and equilibrium quantity in a market

Now that demand and supply have been brought together, it is possible to consider the effects of changes in demand and supply on **equilibrium price** and **equilibrium quantity**.

In Figure 2.16, there has been an increase in the demand for a product — for example, as a result of an increase in incomes in an economy. The demand curve shifts to the right and there is a movement along the supply curve. Equilibrium price goes up from *P*0 to *P*1 and equilibrium quantity increases from *Q*0 to *Q*1.

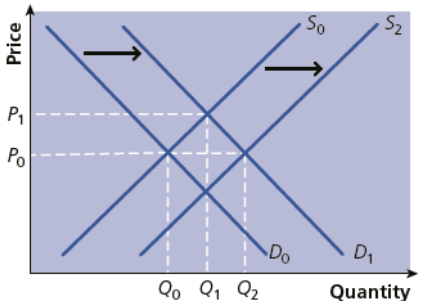


F2.17 The effect of a shift of a supply curve to the right on equilibrium price and equilibrium quantity in a market

In Figure 2.17, there has been an increase in the supply of a product — for example, as a result of a reduction in the costs of production. The supply curve shifts to the right and there is a movement along the demand curve. Equilibrium price falls from *P*0 to *P*1 and equilibrium quantity increases from *Q*0 to *Q*1.

Application of demand and supply analysis

Demand and supply analysis can be applied to a wide variety of different situations. For example, if an economy is experiencing an increase in incomes, there is likely to be an increase in the demand for cars, shifting the demand curve for cars to the right. At the same time, an improvement in technology may have reduced the cost of producing cars, shifting the supply curve to the right. The effect of these two changes:



F2.18 An application of demand and supply analysis to cars

» The demand curve shifts to the right; the effect of this is that equilibrium price rises from P0 to P1 and equilibrium quantity increases from Q0 to Q1.

» The supply curve also shifts to the right; the effect of this is that equilibrium price falls back down to P0 and equilibrium quantity increases from *Q*1 to *Q*2.

» Of course, whether equilibrium price actually returns to its original position will depend on the extent of the shifts of the demand and supply curves.

KEY SKILLS

Application: demand and supply analysis can be applied to many different situations, such as in relation to cars, houses, electrical products, clothing and holidays.

Diagrams: when drawing a diagram showing shifts of demand and/or supply, you need to make sure that the shifts are in the right direction. It is useful to include arrows in the diagram to show clearly the direction of the shifts. It is also important that you clearly label the points of equilibrium in relation to both price and quantity.

The relationships between different markets

- **joint demand:** a situation where two items are consumed together (i.e. they are complements); an example is shoes and shoe laces
- **alternative demand:** a situation where two items are substitutes (i.e. one will be consumed or the other); an example is tea and coffee
- **derived demand:** where demand for the components of a product or for workers arises from demand for the final product

Joint demand (complements)

The existence of complements gives rise to the concept of **joint demand** where goods are consumed together. It would be expected that an increase in the sales of one would lead to an increase in the sales of the other.

Alternative demand (substitutes)

Some goods, on the other hand, are seen as examples of alternative demand: that is, they are in competition with each other and either one is demanded or the other. It would be expected that an increase in the sales of one would lead to a decrease in the sales of the other. Goods that are in **alternative demand** are known as ‘substitutes’.

Derived demand

Derived demand is where the demand for a component depends upon the final demand for a product that uses that component. Derived demand can also be used in relation to the demand for workers.

KEY SKILLS

Application: another example of joint demand where two complementary goods are consumed together is printers and ink cartridges.

Application: another example of alternative demand, where there are two substitute products in competition with each other, is CDs and vinyl records.

Application: derived demand stresses the link between the demand for a component and the demand for a final product that requires that component (e.g. the demand for rubber is derived from the demand for car tyres). The concept can also be applied to the demand for workers (e.g. the demand for bus drivers derives from people’s demand for bus transport).

Joint supply

- **joint supply:** a situation where the process of producing one product leads to the production of another product

Joint supply occurs when the production of one good involves the production of another.

It can often take place in the chemical industry where one chemical may be produced as a by-product of the production of another. It would be expected that a fall in the market price of one may affect the quantity supplied of the other. However, an increase in the price of one of the products could mean that a firm will decide to produce more of both goods.

KEY SKILL

Application: meat and leather is an example of joint supply when the production of one good involves the production of another.

The functions of price in resource allocation

Role of prices

- **rationing:** the rationing function of the price mechanism occurs where demand exceeds supply, leading to a rise in price, a greater scarcity of resources and therefore a rationing of these resources
- **signalling:** the signalling function of the price mechanism occurs where changes in price provide information to both individuals and firms about changes in market conditions, with prices rising and falling to signal scarcities and surpluses in the market
- **price mechanism:** the operation of changes in prices in a market to act as signals to producers to allocate resources according to changes in consumer demand
- **transmission of preferences:** the willingness (or not) of consumers to pay particular prices and so indicate their choices, sending information to producers about their preferences in relation to changing needs and wants
- **incentivisation:** the incentivisation function of the price mechanism is where individuals or firms are encouraged to act in a certain way as a result of higher or lower prices in a market (e.g. individuals are incentivised by the prospect of greater satisfaction, while firms are incentivised by the prospect of greater reward in the form of profit)

Prices perform three important roles in the allocation of resources in a market:

1. **Rationing:** prices perform an important function in a market as a rationing mechanism. If a producer has a limited capacity to produce certain products, when these products are expensive it will have the effect of rationing demand. For example, in the case of exclusive brands of cars, which tend to be very expensive, the high price will limit demand to only those people who can afford to pay this high price.

2. **Signalling:** the **price mechanism** allocates resources because price changes act as signals when the conditions of demand and supply in a market change. The Scottish economist Adam Smith (1723–90) argued that prices in a market therefore act as an 'invisible hand' in allocating scarce resources. This signalling function of the price mechanism is very important in the **transmission of preferences** — it is the way in which consumers indicate their preferences for one product rather than another.

3. **Incentivisation:** for a competitive market to work efficiently, all economic agents (e.g. individuals and firms) need to be able to respond to incentives. For example, if there is an increase in the demand for a product and the price goes up, firms know that for selling each product they will receive a higher average level of revenue per unit, and the possibility of earning a higher profit acts as an incentive.

KEY SKILL

Evaluation: you need to be able to offer a judgement about the role of price in the process of resource allocation in an economy, contrasting its benefits and limitations in carrying out this role. For example, one benefit is that the price mechanism acts as a rationing device as a result of consumers expressing their preferences. However, one limitation is that some people have more income and wealth than others, so they have more power to influence the allocation of resources in an economy.

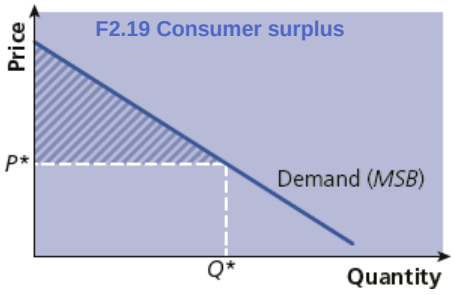
🔗 **Scarcity and choice:** Price plays a key role in enabling consumers to exercise a choice between different products, transmitting a signal to producers in relation to the products that they should produce and therefore how scarce resources will be allocated.

2.5 Consumer and producer surplus

The meaning and significance of consumer surplus

- **consumer surplus:** some consumers will value a particular product more highly than other consumers and yet they will pay exactly the same price for it as the other consumers; this extra satisfaction is consumer surplus and is shown on a demand and supply diagram by the triangle between the price line and the demand curve. It refers to the difference between what a consumer is willing to pay and what they are actually required to pay.

» Consumers are able to obtain a value from consuming a particular product that is above the price paid until at some point consumers pay a price that is exactly equal to the value gained.



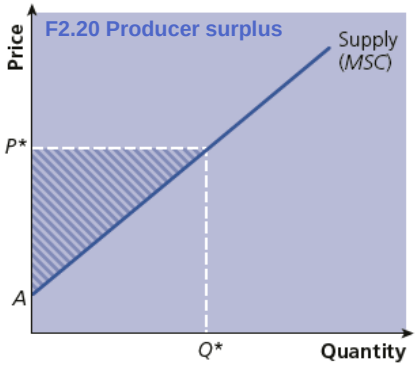
» In the diagram, this is *P**. All the consumers up to *Q** have gained a value that is above the price and this is shown by the shaded area between the price line and the demand curve. When price is *P** and quantity is *Q**, the consumer surplus has disappeared.

» The demand curve is actually showing the marginal social benefit (*MSB*) of the consumption. This means that the demand curve combines all the points where consumers are gaining from the fact that one price is being charged to all consumers in the market, despite the fact that they would have been prepared to pay more (e.g. for concert tickets). They are gaining a marginal social benefit by being able to buy the product at a lower price than they were originally prepared to pay.

The meaning and significance of producer surplus

- **producer surplus:** the difference between the price that consumers are willing to pay for a particular product and the price that producers require in order to supply it; it is shown on a demand and supply diagram by the triangle between the price line and the supply curve

» Producers are able to gain because for all the units sold up to *Q**, they receive a price that is above the cost of producing those units.



» The supply curve actually shows the marginal social cost (*MSC*) of the production. This means that a firm will gain because the price charged is higher than the cost of production, as shown by the supply curve.

» In Figure 2.20 the producer surplus is shown by the shaded area between the price line and the supply curve. When price reaches *P** and quantity is *Q**, the producer surplus has disappeared.

Causes of changes in consumer surplus and producer surplus

When there is a change in the market price, the consumer surplus will also change.
» For example, if the equilibrium price increases above *P** in Figure 2.19, and equilibrium quantity falls below *Q**, the extent of the consumer surplus will be reduced. This is because some consumers will be unwilling to pay the higher price.

» On the other hand, a fall in the market price below *P** would lead to an increase in the consumer surplus. This is because consumers will now be paying less for the product.

When there is a change in the quantity supplied of a product, the producer surplus will also change.

» For example, if the equilibrium price increases above *P** in Figure 2.20, and equilibrium quantity increases above *Q**, the extent of the producer surplus will increase as a result of the rise in the quantity supplied.
» In contrast, if the quantity supplied falls, the extent of the producer surplus will be less. For example, if the equilibrium price falls below P* in Figure 2.20 and the equilibrium quantity decreases below Q*, there will be less producer surplus.

The significance of price elasticity of demand and price elasticity of supply in determining the extent of changes in consumer surplus and producer surplus

Price elasticity of demand is significant in influencing the size of the consumer surplus:

» When the price elasticity of demand is inelastic, consumer surplus will be relatively large.
» When the price elasticity of demand is elastic, consumer surplus will be relatively small.
» When the demand for a product is perfectly elastic, consumer surplus is zero because the demand curve is a horizontal line and the price that consumers pay matches exactly what they are willing to pay.
» On the other hand, when demand is perfectly inelastic, consumer surplus is infinite. In this situation, demand does not respond to a price change.

Price elasticity of supply is significant in influencing the size of the producer surplus:

» Price elasticity of supply is inversely related to producer surplus. If price elasticity of supply is perfectly elastic, the supply curve will be a horizontal line and producer surplus will be zero.
» On the other hand, when supply is perfectly inelastic, it is shown as a vertical line and producer surplus is infinite.

KEY SKILL

Evaluation: you need to be able to make a judgement about the potential impact of both price elasticity of demand and price elasticity of supply in determining the extent of changes in consumer surplus and producer surplus. For example, when demand is price elastic, the impact on the consumer surplus will be greater because consumers will react more strongly than if demand is price inelastic. When supply is price inelastic, and firms are unable to change their output significantly, the impact on the producer surplus will be relatively small.

🔗 **Efficiency and inefficiency:** individual markets, and the economy as a whole, can be both efficient and inefficient in different ways when using scarce resources. Consumer surplus and producer surplus are examples of this. Economic efficiency is where it is impossible to improve the situation of one party without imposing a cost on another. Economic surplus is the sum of the consumer surplus and the producer surplus and this is larger at the equilibrium price and equilibrium quantity than it will be at any other price and quantity. However, if a deadweight loss exists in a market, this is an indication of inefficiency.

3 Government microeconomic intervention

3.1 Reasons for government intervention in markets

Addressing the non-provision of public goods

» Public goods were discussed in section 1.6. It was pointed out there that it would not be possible to provide a public good through a market because it would be impossible to prevent someone who had not paid from benefiting from the service.
» This is known as the 'free rider problem' and comes about as a result of public goods having the characteristics of non-rivalry and non-excludability.
» Therefore, a major reason for government intervention in a market is to address the problem of the non-provision of public goods, an example of market failure because of the impossibility of charging a price for them.

KEY SKILLS

Application: examples of public goods include street lighting, police and defence.

Analysis: you need to be able to explain why government intervention is necessary in relation to the provision of public goods as a result of the impossibility of charging a price for such a product.

Addressing the overconsumption of demerit goods and the underconsumption of merit goods

Merit goods

» Another reason for government intervention in markets is to address the problem of the underconsumption of merit goods.
» Merit goods were discussed in section 1.6. It was pointed out there that merit goods, such as education and healthcare, would be under consumed in a market as a result of imperfect information.
» The problem is that there is information failure and people do not fully appreciate the value of a merit good.
» Therefore, a major reason for government intervention in a market is to address the problem of market failure in the form of the underconsumption of merit goods.

Demerit goods

» Governments also intervene in markets to address the problem of the overconsumption of demerit goods.
» Demerit goods were discussed in section 1.6. It was pointed out there that demerit goods, such as alcohol and tobacco, would be overconsumed in a market as a result of imperfect information.
» The problem is again that there is information failure and people do not fully understand the potential dangers associated with the consumption of products such as alcohol and tobacco.
» Therefore, a major reason for government intervention in a market is to address the problem of market failure in the form of the overconsumption of demerit goods.

KEY SKILL

Analysis: the analysis of government intervention in a market to encourage the consumption of merit goods and discourage the consumption of demerit goods needs to be based on the existence of information failure. This lack of information means that merit goods are undervalued while the potential dangers of demerit goods are not fully understood.

Controlling prices in markets

Another reason for government intervention in markets is to control prices. This intervention can be due to three possible situations:

» **High prices:** the price of certain essential goods in a market, such as bread or rice, could rise so high, without maximum price controls, that poorer sections of a community would not be able to afford them and this could have detrimental effects on their health and standard of living. A government might therefore decide to intervene in the market to prevent the price from rising above a certain level (see section 3.2 of this chapter).
» **Low prices:** the price of certain goods in a market could fall so low, without minimum price controls, that certain producers could go out of business. One example would be intervention in certain agricultural markets to help producers maintain their incomes. A minimum price could also be established in a market for demerit goods, such as tobacco and alcohol, to discourage consumption (see section 3.2 of this chapter).
» **Unstable prices:** if left to free market forces, there is always a chance that prices will fluctuate widely in those markets where there can be great variations in supply over a period of time due to the weather. This is especially the case with agricultural markets, where supply is relatively fixed in the short run, and in such a situation a government might need to intervene in the market.

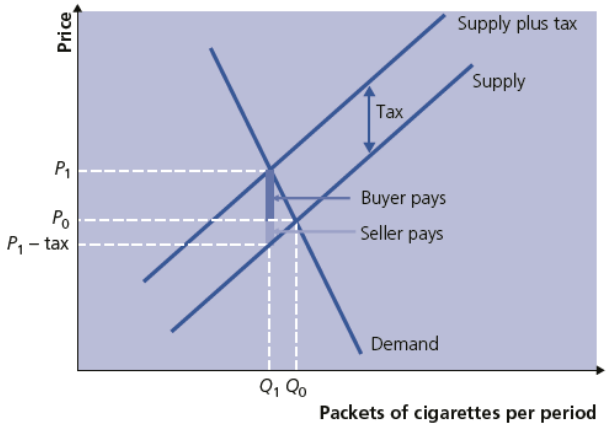
3.2 Methods and effects of government intervention in markets

The impact and incidence of specific indirect taxes

- **indirect tax:** a tax that is imposed on expenditure; it is indirect in that the tax is only paid when the product on which the tax is levied is purchased
- **excise duty:** an indirect tax on expenditure by consumers on such products as fuel, alcohol and tobacco

In the case of a demerit good, a government could intervene in a market through taxation in an attempt to discourage consumption of the good.
» For example, if an **indirect tax**, such as **excise duty**, is placed on demerit goods to such an extent that the price is substantially increased, this is likely to discourage the level of consumption of the demerit good.
» The equilibrium in a market without government intervention, where the demand and supply curves intersect, would be a price of *P*0 and a quantity of *Q*0.
» A government, however, decides to intervene by imposing an indirect tax so that the supply curve shifts upwards by the extent of the tax to 'Supply plus tax'.
This increases the price to *P*1 with the effect that the quantity demanded decreases to *Q*1.
» The diagram shows how the incidence of the tax falls partly on the seller, but mainly on the buyer.

F3.1 The effects of an indirect tax on cigarettes



KEY SKILL

Analysis: the extent to which an indirect tax on a demerit good would discourage consumption depends on the price elasticity of demand for the product. The *PED* for many demerit goods is relatively inelastic, making it less likely that an indirect tax on a demerit good would be very effective in discouraging the level of consumption for such a product.

Distinguishing between impact and incidence

- **impact of tax:** the person, company or transaction on which a tax is levied
- **incidence of tax:** how the burden of taxation is shared between the producer and consumer
- **specific tax:** an indirect tax that is a fixed amount per unit of output

It is important to distinguish between the impact and the incidence of specific indirect taxes.

The **impact of a tax** refers to the company on which, or the person on whom, a tax is levied: that is, the entity legally responsible for handing the tax over to the authorities.

The impact of a tax, therefore, is essentially concerned with the legal situation, i.e. who has to pay the tax to a government.

The **incidence of a tax**, however, refers to where the eventual burden of the tax falls — that is, how the payment of a tax is divided between different people:

» For example, with an indirect tax on a retailer, the burden of the tax is likely to be shared between the producer and the consumer.
» The more inelastic is the demand, and the more elastic is the supply, the greater the burden will be on the consumer. If the price elasticity of demand is perfectly inelastic, and there is a vertical demand curve, the incidence of the tax will be entirely on the consumer.
» On the other hand, if the price elasticity of demand is perfectly elastic, and there is a horizontal demand curve,
the incidence of the tax will be entirely on the producer.

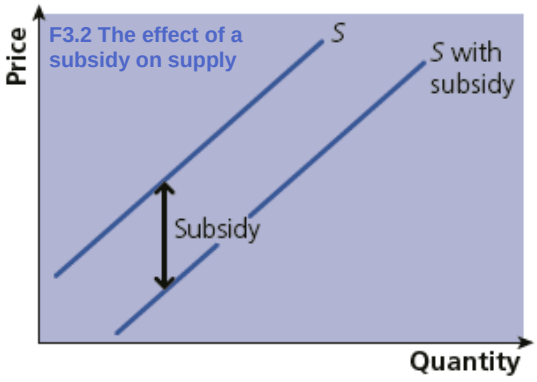
Indirect taxes on expenditure can take different forms

» An excise duty, for example, is usually a **specific tax** on a product — in other words, a specific amount is required to be paid, not a percentage of the selling price.
» An ad valorem tax, on the other hand, requires a percentage of the selling price to be paid.

The impact and incidence of subsidies

- **subsidy:** an amount of money paid by a government to a producer, so that the price charged to the customer will be lower than would have been the case without the subsidy
- **impact of a subsidy:** the effect of a subsidy on the price and quantity of a product in a market
- **incidence of a subsidy:** how the gain from a subsidy is shared between the producer and consumer

» **Subsidies** are an example of microeconomic government intervention to encourage the production and consumption of a particular product.
» If a government pays firms a subsidy to produce a particular product, this will reduce their costs and encourage firms
to supply more output at any given price. In Figure 3.2 the supply curve shifts downwards to the right.



» The result of the subsidy, in shifting the supply curve to the right, will be a lowering of price. The actual price will be determined where the '*S* with subsidy' line intersects with the demand curve. Both producers and consumers may therefore benefit from the subsidy.

» The **impact of a subsidy** in a market will be a reduction in price and an increase in output. The **incidence of a subsidy** relates to who is made better off by the subsidy and by how much.
» The vertical distance between the two supply curves in Figure 3.2 indicates the size of the subsidy, but the price that the consumer pays does not fall by the full amount of the subsidy. This is because the producer gains some of the benefit in terms of extra revenue that they can keep.
» The effect of the subsidy is that there will be both a gain to the consumer and a gain to the producer. The extent of the different gains will depend on the price elasticity of demand for, and the price elasticity of supply of, the product.

The direct provision of goods and services

- **direct provision of goods and services:** where a government decides to provide particular goods and services itself

Although a government can intervene in a market, such as through indirect taxes and subsidies, it is also possible for a government to provide goods and services directly.

» A government could establish the **direct provision of goods and services** alongside the private sector, financed by the revenue received from taxation. This is likely to be the case with certain merit goods, such as the provision of healthcare and education. In many countries, these services are provided through both the public and private sectors.
» Of course, a government may decide to provide a good or service through the public sector by *nationalising* an industry (i.e. taking it under state control).

Maximum and minimum prices

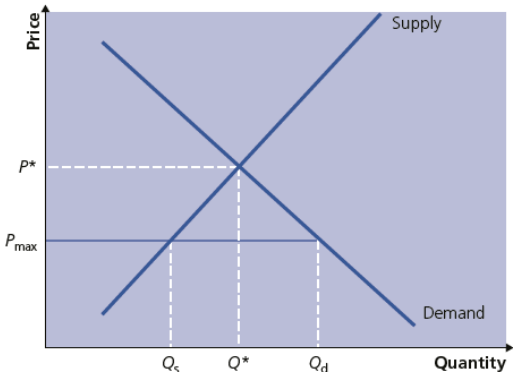
Maximum price controls

maximum price: a price that is fixed in a market and which the price must not exceed

One form of government intervention in an economy is to establish a **maximum price** in a market for a product — a level above which the price cannot rise.

» This maximum price needs to be set below the equilibrium price that would have resulted from the intersection of demand and supply.

» Figure 3.3 shows that *P** and *Q** would be the equilibrium price and equilibrium quantity in a market without government intervention.



F3.3 The effect of a maximum price in a market

» However, a government could decide to establish a maximum price below *P** at *P*max. The quantity demanded would now be *Q*d, but the quantity supplied would be *Q*s. There would therefore be excess demand in the market, creating a shortage, and this could lead to queuing, rationing or the emergence of a black market.

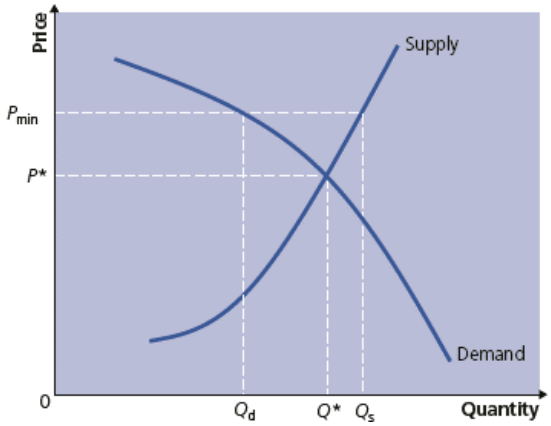
KEY SKILL

Application: examples of goods that could benefit from the establishment of a maximum price in a market include essential foods (e.g. bread and rice). This is because without maximum price controls, the price of these foods could rise so high that poorer sections of a community would not be able to afford them. This could have detrimental effects on their health and standard of living.

Minimum price controls

- **minimum price:** a price that is fixed in a market and below which the price must not fall

The establishment of a maximum price in a market for a product is designed to prevent the price rising above a specific level. However, it is also possible that a government will wish to intervene in a market to set a **minimum price** to prevent the price falling below a specific level.



F3.4 The effect of a minimum price in a market

» The minimum price will need to be established above the equilibrium price that would have resulted from the intersection of demand and supply.
» Figure 3.4 shows that P^* and Q^* would be the equilibrium price and equilibrium quantity in a market without government intervention.

» However, a government could decide to establish a minimum price above P^* at P_{min}. The quantity demanded would now be Q_d, but the quantity supplied would be Q_s.
» There would therefore be excess supply in the market, creating a surplus, and this could lead to producers becoming inefficient and the emergence of a black market where demerit goods, such as tobacco and alcohol, are sold at prices below the minimum price.

KEY SKILL

Application: examples of minimum price controls include intervention in certain agricultural markets to help producers maintain their incomes.

Buffer stock schemes

- **price stabilisation:** where a government intervenes to purchase stocks of a product when supply is high and to sell socks of a product when supply is low
- **buffer stock:** a stock of a commodity that is held back from the market in times of high production and released onto the market in times of low production

If a market is characterised by wide fluctuations in prices, such as agricultural markets, a government could try to bring about **price stabilisation** by intervening through what is called a **buffer stock scheme**:

» A target price is decided.

» When there is a good harvest, supply is very high in the market and the government purchases some of the stock and stops it from entering the market; the effect of this is to stop the price going too low.

» When there is a bad harvest, supply is very low in the market and the government releases some of this stock; the effect of this is to stop the price going too high.

🔗 **Efficiency and inefficiency:** a buffer stock scheme could encourage the oversupply of products if producers know that the government will buy any excess supply. So, the scheme could encourage producers to be less efficient.

The advantages and disadvantages of a buffer stock scheme:

Advantages	Disadvantages
• The scheme provides greater stability of prices, which could encourage more investment in agriculture. • It ensures food supplies and avoids shortages. • It prevents producers going out of business when there is a large drop in prices. • It helps to maintain the incomes of producers.	• The cost of buying the excess supply could be high, which could be a problem for a government that might need to raise taxation to obtain the necessary funds to operate the buffer stock scheme. • The scheme could encourage oversupply if producers know that the government will buy any excess supply. • The scheme could encourage producers to be less efficient. • It may not be possible to store all products in a buffer stock as some may be perishable. • It may be difficult for the government to decide on the target price.

KEY SKILL

Problem solving: some markets, particularly agricultural markets, can experience a great deal of price instability, so a buffer stock scheme can be used to reduce the extent of this volatility in prices.

Provision of information

» Market failure can be caused by inadequate information. A government could therefore aim to increase the availability of appropriate information to consumers in order to try to influence their economic behaviour.
» It is assumed that consumers will always aim to maximise their utility or satisfaction, but they will only be able to achieve this objective if they are in possession of the necessary information. If this information is not available, it is unlikely that they will be able to make rational decisions.
» Information failure is a major cause of market failure, so a government will need to take measures to improve the accuracy and availability of information that consumers need. This will help them make rational decisions and ensure that scarce resources are allocated as efficiently as possible.
» For example, a government could aim to make people as well informed as possible about the potential advantages of consuming merit goods, such as education and healthcare. It could also aim to make people as well informed as possible about the potential disadvantages of demerit goods, such as alcohol and tobacco. A particular example of such an approach in relation to the consumption of demerit goods is nudge theory.

KEY SKILL

Application: an example of nudge theory in relation to demerit goods is when a government puts a health warning on cigarettes, such as ‘Smoking can kill’.

3.3 Addressing income and wealth inequality

The difference between income as a flow concept and wealth as a stock concept

- **income:** money received, especially on a regular basis
- **wealth:** all assets that have a monetary value

It is important to distinguish between **income** and **wealth**. There are clear differences between them and one of the most important is the idea of income as a flow and wealth as a stock.

Income

Income is a flow of money received by factors of production, including:

» **wages and salaries:** these are paid to people for the work they have carried out
» **welfare benefits:** money paid to people in the form of a state pension or a tax credit
» **profits:** these are received by businesses
» **dividends:** payments distributed to shareholders
» **rental income:** this is a flow to people who own, and rent or lease out, property
» **interest:** money paid to people who hold funds in interest-paying accounts with financial institutions

Wealth

Wealth is a stock of money or assets, including:

» **savings:** these can be held in various forms of accounts
» **shares:** ownership of shares issued by limited companies
» **property:** ownership of houses and apartments
» **bonds:** money held in instruments of indebtedness
» **pension schemes:** wealth held in occupational pension schemes and life assurance schemes

🔗 **The role of government and the issues of equality and equity:** a government can intervene in an economy to bring about greater equality and equity in relation to the distribution of income and wealth.

Measuring income and wealth inequality

The Gini coefficient

- **Gini coefficient:** a statistical measurement of the degree of inequality of income in an economy

A number of policies could be used with the objective of bringing about a redistribution of income and wealth in an economy. However, it is necessary first to consider how that distribution can be measured. This is achieved using a **Gini coefficient:**

» A Gini coefficient is a way of measuring the extent of inequality in the distribution of income in an economy. It is measured by the ratio of the area between the diagonal line of total equality and the Lorenz curve, on the one hand, to the total area under the diagonal line, on the other (see Chapter 11, section 11.4, for an explanation of the Lorenz curve).
» In this way, the Gini coefficient measures the extent to which the distribution of income in an economy diverges from the position of total equality. The lower the value of the coefficient (i.e. the closer the Gini coefficient is to 0), the more even is the distribution of income. The higher the value of the coefficient (i.e. the closer the Gini coefficient is to 100), the less even is the distribution of income.

Economic reasons for inequality of income and wealth

There are a number of possible economic reasons for the inequality of income and wealth, including the following:

» **Employment:** a major cause of income equality is variations in the ability of people to access well-paid employment. When there is an increase in unemployment in an economy, there will be fewer people receiving wages and salaries and more people receiving benefits. When there has been a decrease in full-time employment and an increase in part-time employment, income inequality will widen because rates of pay are usually lower in part-time than in full-time employment.
» **Government policy:** for many workers, a difficult economic situation could have given rise to a 'wage freeze' or lower-than-inflation wage rises. This is particularly true of public sector workers who may have experienced a fall in their real standard of living when increases in wages have been less than increases in prices.
» **Taxation:** a government may decide to raise the levels of taxation in order to increase public revenue. It might also make tax more regressive, taking proportionally more tax from people on lower incomes than those on higher incomes.
» **Distribution of wealth:** people who already hold wealth are able to invest, which creates new wealth. The existing concentration of wealth makes inequality a vicious cycle.

Policies to redistribute income and wealth

Minimum wage

- **minimum wage:** the lowest wage permitted by law (i.e. a price floor below which employers cannot pay their employees)

A **minimum wage** is one way to redistribute income and wealth in an economy:

» A government could decide to establish a minimum wage rather than allow wages to be determined by the forces of the demand for, and the supply of, labour in particular markets.
» One problem with a minimum wage, however, is that it could lead to income becoming more unequal. For example, some employers may not be able to pay all their workers the minimum wage, so some of them may become unemployed and be forced to live on benefits.
» Another potential problem is that a national minimum wage takes no account of variations in the cost of living in different parts of a country.
» In some countries, there may be a large informal economy and a minimum wage would not apply to such workers.

🔗 **Problem solving:** a major problem in all economies is inequality of income and wealth. Governments can use policies in an attempt to redistribute income and wealth. These include a minimum wage, transfer payments, progressive taxes and state provision of essential goods and services.

Transfer payments

- **transfer payment:** a form of payment to those in society who are less well off, paid for out of the revenue received from taxation

A government could decide to intervene in a market through the use of **transfer payments**:

» This means that revenue received from taxation is used to give financial support to people, such as in the form of pensions and benefits.
» These transfer payments can be regarded as worthwhile because they are made to those in society who are less well off.

» Transfer payments can be criticised, however, for having a distorting effect: for example, in some countries, people who are unemployed can be given financial support in the form of a benefit, but if this benefit is too high, it may make such people less inclined to look for work.

Progressive income taxes, inheritance and capital taxes

Income taxes

» A government could address income inequality through the use of progressive taxes.
» Many economies use progressive taxation to achieve the macroeconomic objective of a fairer and more equitable distribution of income.
» An income tax, for example, will not only take more from a person as their income rises, but a higher *proportion* of that income.

Inheritance and capital taxes

» A government may also decide to intervene in an economy to try to bring about a more equitable distribution of wealth, as well as income.
» Examples of taxes that can achieve this objective are progressive inheritance and capital taxes.
» Inheritance taxes apply to situations where money and/or property have been passed on to someone in the event of a death, while capital taxes apply to situations where a profit has been made by the selling of certain assets such as paintings or antiques.

State provision of essential goods and services

» State provision of essential goods and services helps to redistribute income and wealth because the money to pay for the provision of these essential goods and services comes from the money received from taxation.
» This helps to lessen inequality because the money received by the government pays for the goods and services that poorer people are unable to afford. State provision of healthcare is a good example of this.
» However, state provision of essential goods and services can be criticised for a number of reasons. For example, the goods and services might be provided to some people who would have been able to pay for them, so some public money could be wasted. Also, the quantity and quality of provision would be limited by the amount of funds available to the government and this would depend on the economy's fiscal policy and the amount of tax revenue collected.

4 The macroeconomy

4.1 National income statistics

The meaning of national income

- **national income:** a general term for the total income of an economy over a particular period of time

National income is often used as a generic term, but there are actually three different forms of national income statistics:

» gross domestic product (GDP)
» gross national income (GNI)
» net national income (NNI)

The measurement of national income

- **gross domestic product (GDP):** the total value of all that has been produced over a given period of time within the geographical boundaries of a country
- **gross national income (GNI):** the gross domestic product of a country plus net income from abroad
- **net national income (NNI):** the gross domestic product of a country plus net income from abroad minus the depreciation of fixed assets

Gross domestic product

Gross domestic product (GDP) refers to all that is produced within the geographical boundaries of a particular country. It does not matter whether the productive assets are owned locally or foreign owned.

There are three different ways of measuring the value of a country's GDP:

» **The output method:** this adds up the total amount of output produced by firms in an economy — in terms of the value added by each firm to avoid double counting.
» **The income method:** this adds up the total amount of income received by people in an economy, such as in terms of wages, salaries and profits.
» **The expenditure method:** this adds up all of the spending in an economy by households, firms and the government, including net export spending.

Each of the three approaches will produce the same figure because they all measure the flow of income in an economy over a particular period of time.

Gross national income

Gross national income (GNI) is GDP plus net income from abroad.

Net national income

Net national income (NNI) is calculated by taking GDP and adding to it the net receipts of wages, salaries and property income from abroad, minus the depreciation of fixed capital assets.

The adjustment of measures from market prices to basic prices

GDP deflator

- **at current market prices:** data that are expressed in terms of the prices of a particular year (i.e. they have not been adjusted to take account of inflation)
- **at constant prices:** data that have been adjusted to take into account the effects of inflation
- **GDP deflator:** a ratio of price indices that is used in national income statistics to remove the effect of price changes, so that the figures can be seen as representing real changes in output
- **exports:** goods and/ or services that are produced domestically in one country and sold to other countries
- **imports:** goods and/ or services that are produced in foreign countries and consumed by people in the domestic economy

» It is important to distinguish between nominal value and real value.
» If a country is experiencing inflation, the value of its GDP will rise, but this increase could be due solely to the rise in prices — in other words, there may not have been a real increase in value if the effect of inflation is eliminated from the figures. This is a limitation of any data expressed **at current market prices**. It is therefore important to distinguish between nominal and real variables.

» Economists usually produce national income statistics **at constant prices**, so that changes in real output can be identified rather than changes in value that are purely due to the inflation that exists in a country.
» The **GDP deflator** is a price index that is used to convert the figures into real GDP. It measures the prices of products produced in a country and not the prices of products consumed.
» It therefore includes the value not just of consumer products but also of the capital used in the production of the products. It includes the prices of **exports**, but not the prices of **imports**.

KEY SKILL

Interpretation of data: it is important to consider whether data are nominal or real. Nominal data have not been adjusted to take into account the effect of inflation, whereas real data have been adjusted to take inflation into account.

🔗 It is important to demonstrate a clear understanding of the difference between a real change in the value of a country's output, after taking into account the effects of inflation, and a change in the value of a country's output that is purely due to inflation in that economy.

The adjustment of measures from gross values to net values

- **depreciation (of capital):** the decline in the value of a capital asset over a given period of time (usually a year)
- **net domestic product (NDP):** the gross domestic product of a country minus depreciation or capital consumption
- **net national product (NNP):** the gross national product of a country minus depreciation or capital consumption

KEY SKILL

Interpretation of data: it is important to consider whether data have been adjusted to take into account the effect of depreciation. If this has not happened, the data are gross; if it has happened, the data are net.

🔗 Gross values are adjusted to net values by the deduction of **depreciation** or capital consumption. This refers to where capital equipment has worn out or broken down.

Net domestic product

Net domestic product (NDP) is obtained by deducting depreciation from GDP.

Net national product

Net national product (NNP) is calculated by deducting depreciation from the gross national product. As with net domestic product, this is done to take into account the money that will need to be spent on replacing machinery and equipment that has worn out during the course of the year.

4.2 Introduction to the circular flow of income

The circular flow of income in a closed economy and an open economy

- **circular flow of income:** the flow of income around an economy, involving a mixture of injections and withdrawals or leakages
- **closed economy:** an economy that does not trade with the rest of the world
- **open economy:** an economy that trades with the rest of the world

The **circular flow of income** refers to the flow of income and spending around an economy. It is a model of the economy in which the major exchanges are represented as flows of money between different economic agents.

It is important to distinguish between a closed economy and an open economy:

» **Closed economy**: a basic approach is to consider the circular flow of income in a closed economy. In a closed economy, it is assumed that a country does not trade with any other countries. If the circular flow of income is limited to the movement of incomes between households and firms, it is known as a 'two-sector economy'. If government is then added to the circular flow, it becomes a 'three-sector economy'.

» **Open economy**: a more realistic approach would be to consider the circular flow of income in an open economy. In an open economy, it is assumed that a country does trade with other countries and so flows of money include the goods and services that are exported to, and imported from, other countries.

In an open economy, therefore, the circular flow of income involves four groups:

» households
» firms
» the government
» the international economy

This is why such an economy is known as a **'four-sector economy'**.

KEY SKILL

Analysis: it is necessary to include the impact of all four sectors (i.e. households, firms, government and the international economy) when analysing a four-sector economy. This is because a four-sector economy includes all possible aspects of economic activity.

Injections and leakages

injection: spending that adds to the circular flow of income; this can come from investment, government expenditure and exports
leakage (or withdrawal): money that leaks out of the circular flow of income; this can be as a result of savings, taxation and imports

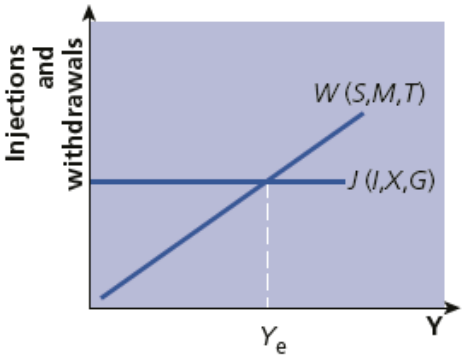
At any one time, there will be a number of **injections** of money into an economy and a number of **withdrawals** or **leakages** of money out of an economy.

The basis of aggregate demand in an economy is consumption expenditure by households. Then the injections can be added to, and the withdrawals or leakages taken from, this expenditure, as shown below:

Injections	Withdrawals or leakages
Investment spending by private sector firms (*I*)	Savings (*S*)
Government spending (*G*)	Taxation (*T*)
Income from exports sold abroad (*X*)	Income spent on imports from abroad (*M*)

Candidates need to ensure that they understand, and can refer to the different injections into, and the various leakages/withdrawals from, the circular flow of income.

Equilibrium and disequilibrium



F4.1 Equilibrium using the injection/withdrawal approach

» One way of showing equilibrium in an economy is through the injection/ withdrawal approach.
» For the overall economy to be in equilibrium, the injections into the circular flow of income need to be equal to the withdrawals from the circular flow of income. where the injections (*I*, *G* and *X*) are equal to the withdrawals (*S*, *T* and *M*) at the *Ye* level of income.
» Equilibrium occurs when there is no stimulus to change any of the values in the circular flow of income — that is, equilibrium is defined as a situation in which there is no tendency for the levels of income, expenditure and output to change.
» The condition for equilibrium in the macroeconomy is when total planned injections are equal to total planned withdrawals.
» When injections are more than withdrawals (i.e. the additions from *I*, *G* and *X* are more than the withdrawals from *S*, *T* and *M*), there will be a disequilibrium and this will cause the level of national income to rise.
» If the withdrawals are more than the injections, this will cause the level of national income to fall.

🔗 **Equilibrium and disequilibrium:** individual markets, and the economy as a whole, are always moving into and out of equilibrium, constantly altering the allocation of resources. In this particular context, equilibrium occurs in the macroeconomy when injections equal withdrawals. Whenever they are not equal, disequilibrium will occur.

4.3 Aggregate demand and aggregate supply analysis

The definition of aggregate demand

Aggregate demand *(AD)* refers to the total amount of goods and services demanded in an economy at a given overall price level at a given time.

The components of aggregate demand and their meanings

- **aggregate demand (*AD*):** the total amount that is spent on an economy's goods and services at a given price level over a given period of time; it is made up of four main components: **consumption**, **investment**, **government spending** and **net exports**

There are four main components of aggregate demand:

» consumer spending by households on good and services (*C*)
» investment by firms in machinery and equipment (*I*)
» government spending (*G*)
» the net effect of international trade, i.e. exports (*X*) minus imports (*M*)

Aggregate demand is therefore the sum of consumption expenditure, investment expenditure, government expenditure and net exports. It can be shown as:

$AD = C + I + G + (X - M)$

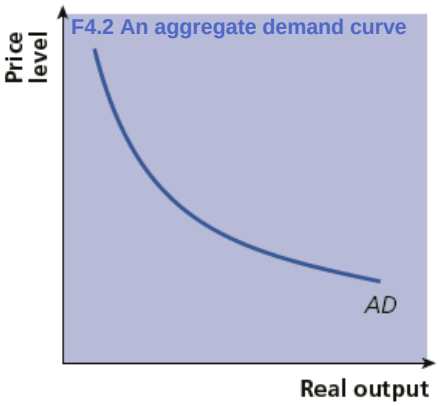
The determinants of aggregate demand

A change in any of the four main components of aggregate demand will bring about a change in aggregate demand. These changes could come about as a result of a number of factors, including but not limited to changes in:

» interest rates (i.e. the price of money)
» the money supply (i.e. the quantity of money)
» taxation
» expectations of future economic conditions
» degrees of confidence (e.g. optimism or pessimism)
» exchange rates
» the accumulation of income/wealth
» technology

The shape of the aggregate demand curve

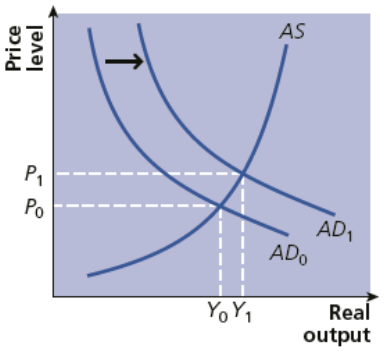
The aggregate demand curve slopes downwards from left to right. This can be seen in Figure 4.2. It shows the relationship between the general price level and real output. If there is a fall in the price level, this will cause a movement along the *AD* cu rve to the right because with goods cheaper, consumers effectively have more spending power. This is known as the 'real money balance effect' and leads to an expansion of aggregate demand. It is assumed that the components of aggregate demand are constant at this time.



F4.2 An aggregate demand curve

The causes of a shift in the aggregate demand curve

» If there is a change in any of the components of *AD*, such as a change in consumption expenditure, investment expenditure, government expenditure or net expenditure on exports, the *AD* curve will shift.
» If there is an increase in aggregate demand, the *AD* curve will shift to the right. This can be seen in Figure 4.3 where the *AD* curve shifts from *AD*0 to *AD*1.
» If there is a decrease in aggregate demand, the *AD* curve will shift to the left.



F4.3 A shift in aggregate demand

The definition of aggregate supply

- **aggregate supply (*AS*):** the total output that firms in an economy are able and willing to supply at different price levels in a given period of time; it includes both consumer and capital products

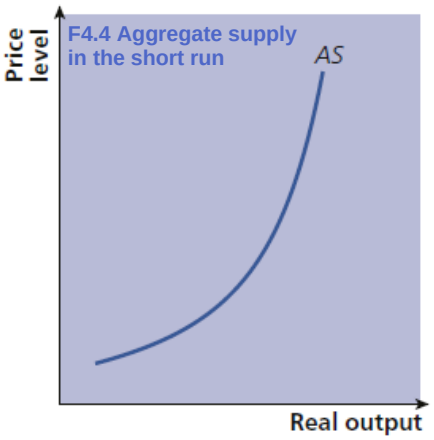
The determinants of aggregate supply

Changes in aggregate supply are caused by a number of factors, including but not limited to changes in the following:

» the state of technology
» the cost and productivity of capital
» the cost and productivity of labour
» the cost of raw materials
» taxation
» exchange rates
» government policy (e.g. rules and regulations)

The shape of the aggregate supply curve in the short run and the long run

» The short-run aggregate supply curve (*SRAC*) slopes upwards from left to right.
» If there is a change in the price level, this will cause a movement along the *AS* curve. The upward slope represents increasing marginal costs with an increase in production.
» When the price level is relatively low, the aggregate supply is low; however, as the price level and potential profits increase, aggregate supply increases, creating the upward-sloping *SRAS* curve.

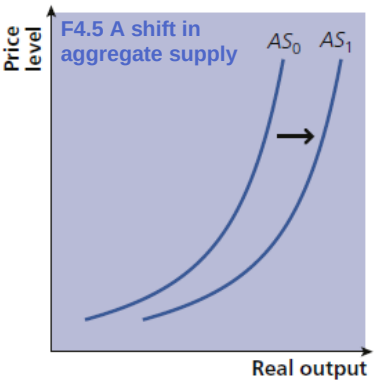


In the long run, however, the shape of the *AS* curve will change:

» At low levels of output, the long-run aggregate supply curve (*LRAS*)
can be horizontal, indicating that it is perfectly elastic.
» At higher levels of output, the *LRAS* curve can be upward sloping.
» At very high levels of output, the *LRAS* curve can be vertical, indicating that it is perfectly inelastic. Indeed, some economists argue that the long-run *AS* curve is perfectly inelastic, indicating that an economy might operate at full capacity.

Causes of a shift in the aggregate supply curve in the short run and the long run

If there is a change for any reason other than a change in the price level, the *AS* curve will shift. If there is an increase in aggregate supply, the *AS* curve will shift to the right. This can be seen in Figure 4.5, where the curve shifts from *AS*0 to *AS*1. If there is a decrease in aggregate supply, the *AS* curve will shift to the left.



The distinction between a movement along and a shift in aggregate demand and aggregate supply

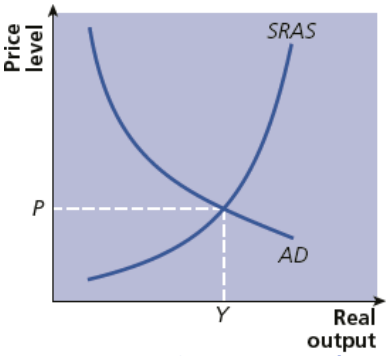
As has already been indicated, it is important to distinguish between a movement along and a shift in *AD* and *AS*.

» Figure 4.2 shows a movement along an *AD* curve as a result of a change in the price level, whereas Figure 4.3 shows a shift in *AD* as a result of a reason other than a change in the price level.
» Figure 4.3 shows the *AD* curve shifting to the right and this could have been caused by an increase in business or consumer confidence or an expansionist fiscal or monetary policy, such as lower tax rates, higher government spending, lower rates of interest or less spending on imports.
» Figure 4.4 shows a movement along an *AS* curve as a result of a change in the price level, whereas Figure 4.5 shows a shift in *AS* as a result of a reason other than a change in the price level.
» Figure 4.5 shows the *AS* curve shifting to the right and this could have been caused by cheaper imported materials, lower money wages, lower rates of interest or improved technology.

The establishment of equilibrium in the *AD*/*AS* model and the determination of the level of real output, the price level and employment

Equilibrium is determined when planned aggregate demand equals planned aggregate supply. This can be seen in F4.6, where the equilibrium price level is *P* and the equilibrium real output level is *Y*.

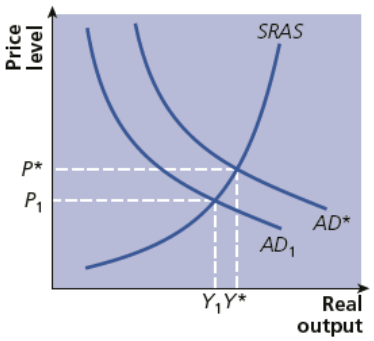
This equilibrium position will not necessarily be at the full employment level for an economy. This can be seen in F4.7. The equilibrium where *AD** crosses *SRAS* (short-run aggregate supply) is at price *P** and real output *Y**. At this point, the macroeconomic equilibrium is at full employment.



F4.6 Macroeconomic equilibrium

However, the equilibrium where *AD*1 crosses *SRAS* is at price *P*1 and real output *Y*1. At this point, the equilibrium is below the full employment level in the economy, so there is surplus capacity.

F4.8 shows the equilibrium position in relation to long-run aggregate supply. As indicated before, aggregate supply is initially elastic, but as it gets nearer to the full employment level, shown by *Y**, it becomes increasingly inelastic, until it eventually becomes perfectly inelastic when the full employment (*Y**) level of output is reached.

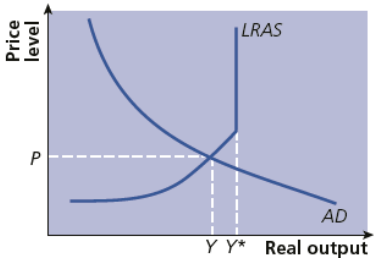


F4.7 Will macroeconomic equilibrium be at full employment?

🔗 **Equilibrium and disequilibrium:** national income equilibrium occurs where *AD* and *AS* intersect. National income disequilibrium occurs where *AD* is no longer exactly matched by *AS*.

🔗 **Time:** time is important in relation to aggregate supply.
In the short run, the *AS* curve will be upward sloping, but in the long run it can become vertical when the full employment level of output has been reached.

🔗 in the Keynesian model of the economy, the equilibrium of *AD* and *AS* will not necessarily be at the full employment level of an economy.



F4.8 Macroeconomic equilibrium revisited

The effects of shifts in the aggregate demand and aggregate supply curves on the level of real output, the price level and employment

» Shifts in the *AD* and/or *AS* curve will have an effect on the level of real output, the price level and employment.
» Figure 4.7 showed the effect of a shift to the left of the *AD* curve from *AD** to *AD*1. The new equilibrium is now where *AD*1 crosses *SRAS* at price *P*1 and real output *Y*1. At this point, the equilibrium shows a fall in real output below the full employment level of output in the economy, so there is surplus capacity. It also shows a lower price level, down from *P** to *P*1.
» Shifts in the *AS* curve will also have an effect on real output, the price level and employment. For example, a shift to the left of the *AS* curve, as a result of more expensive imported materials, higher money wages or higher rates of interest, would lead to an increase in the price level, a decrease in the level of real output and a decrease in employment.

4.4 Economic growth

The meaning of economic growth

- **economic growth:** an increase in the national output of an economy over a period of time, usually measured through changes in gross domestic product

Economic growth is defined as the increase in national output of a country over a period of time. It is possible to distinguish between actual economic growth and potential economic growth.

The measurement of economic growth

Economic growth is usually measured in terms of a change in gross domestic product over a particular period of time, usually one year. It can be shown by an outward shift of a production possibility curve.

KEY SKILL

Numerical skills: you will need to be able to distinguish between the rate of change and the percentage change in relation to economic growth. For example, if GDP in an economy has increased from $300 billion to $320 billion, there has been an increase of $20 billion, and as a percentage this would be an increase of 6.66%.

The distinction between growth in nominal GDP and real GDP

- **nominal GDP:** the value of all the goods and services produced by a country in a given period at current market prices
- **real GDP:** the value of a country's total economic output in a given period, adjusted for the effects of price changes

Economic growth is the increase in the production of goods and services in an economy over a specific period, but it is important to distinguish between nominal and real economic growth. **Nominal GDP** is GDP that has not been adjusted to take into account the effects of inflation, whereas **real GDP** is GDP that has been adjusted to take into account the effects of inflation

The causes of economic growth

- **business (or trade) cycle:** the fluctuations in the national output of a country, involving a succession of stages or phases, including boom, recession, trough and recovery

Economic growth in an economy can be brought about by a number of factors, including:
» An increase in the number of workers
» An improvement in the quality of labour — for example, the acquisition of new skills leading to a higher level of productivity
» a greater commitment to research and development, in terms of both invention (the discovery of new products and new methods of production) and innovation (a process by which a good or service is renewed by applying new processes or introducing new techniques to create new value)
» An improvement in the state of technology
» Investment in capital stock
» a move towards more capital-intensive production
» Increased mobility and flexibility of factors of production
» a more efficient allocation of resources
» Development of new markets for exports
» a reduction in taxes on company profits to allow firms more funds to finance investment
» An upturn in the **business** (or **trade**) **cycle**, which will increase business confidence and lead to firms increasing output

🔗 economic growth is concerned not only with the *quantity* of the factors of production used in the production process, but also with the *quality* of these economic resources. For example, net88 immigration could lead to an increase in the number of workers. However, the productivity of existing workers could be increased as a result of various training and re-skilling schemes.

The consequences of economic growth

The benefits and costs of economic growth:

Benefits	Costs
• Economic growth will lead to an increase in the standard of living of a country, resulting from the greater number of goods and services produced in the economy. • An economy that is growing shows that the economy is doing well; this should result in greater confidence that should encourage future investment. • Economic growth is likely to lead to a decrease in the level of unemployment in the country, although this will depend on the extent to which the extra output is produced through **labour-intensive** or **capital-intensive** methods of production. • An increase in output, much of which is exported to other countries, may lead to the reduction, or possibly the elimination, of a deficit in the balance of payments.	• There can sometimes be a shift away from the production of consumer goods to capital goods in order to bring about economic growth; this could potentially be beneficial in the long run, but not necessarily in the short run. • Economic growth may deplete natural resources and do damage to the environment in terms of various forms of pollution; it is in these senses that a high rate of economic growth is sometimes said to be unsustainable. • The benefits of economic growth may not always be shared evenly among different people in the country. • There may be a reduction in the quality of life in the country (e.g. working hours may be longer, reducing the amount of leisure time).

- **labour-intensive production:** a process of production with a relatively high proportion of labour inputs, compared to other inputs
- **capital-intensive production:** a process of production with a relatively high proportion of capital inputs, compared to other inputs

🔗 If an examination question asks for a discussion of the concept of economic growth, you should remember to include a consideration of both the benefits and the costs of growth for an economy.

🔗 **Progress and development:** economics studies how societies can progress in measurable money terms and develop in a wider, more normative sense regarding living standards, inclusivity and sustainability. Economic growth enables people to have a higher standard of living than would otherwise be the case. However, it must be remembered that although there are a number of potential benefits of economic growth, there are also several potential costs.

KEY SKILL

Evaluation: you will need to be able to evaluate the advantages and disadvantages of economic growth and come to a judgement as to whether the benefits outweigh the costs.

4.5 Unemployment

The meaning of unemployment

- **unemployment:** where a number of people in an economy are able and willing to work but are unable to gain employment

Unemployment refers to the situation which occurs when people are able and willing to work, but are unable to find employment.

Measures of unemployment, with reference to possible difficulties in measurement

- **unemployment rate:** the number of unemployed people divided by the labour force

It is important to distinguish between the number of people who are unemployed in a country and the **unemployment rate**. The unemployment rate refers to the total number of people who are unemployed in a country divided by the labour force and this is expressed as a percentage. Economists are interested in discovering patterns and trends in the rate of unemployment in a country over a period of time. It is useful to establish whether the trend is upward or downward; if it is upward, the government will need to devise appropriate policies to try to reduce the rate.

🔗 Candidates need to distinguish between the *number of people who are unemployed* in an economy and the *rate of unemployment* in the economy. In the first case, it will be a number; in the second case, it will be a percentage.

KEY SKILL

Interpretation of data: it is important to understand not just whether unemployment has increased or decreased over a period of months or years, but whether the overall trend over a period of time has been upward or downward. It is also important to understand that unemployment data can be in the form of absolute values or in percentage terms.

The size and components of the labour force

- **labour force:** the number of people in a country who are employed or who are looking for work

The **labour force** refers to all the people in a country who are employed as well as those who are looking for work. It therefore consists of both the employed people and the people who are unemployed. It is best defined as the number of people in a country who are economically active, either in employment or unemployed.

The size of a country's labour force depends on a number of factors, including:

» The total size of the population
» The birth rate
» The death rate
» The school leaving age
» The number of people who stay in full-time education after leaving school
» The retirement age
» The availability and value of welfare benefits to those who do not have a job
» The availability and cost of childcare
» The attitudes in the society to women working
» The state of the economy, such as whether an economy is in a boom or a recession

🔗 Candidates often seem to believe that the term ‘labour force’ only applies to those people who are employed — that is, to those people who actually have a job. However, it actually refers not only to those people who have a job, but also to those who are unemployed.

The working population

- **working population:** the people in a country who are working or who are actively seeking work
- **dependency ratio:** the number of dependent people in a country divided by those of working age
- **participation rate:** the proportion of the population which is either in employment or officially registered as unemployed
- **claimant count:** the number of people who officially register as unemployed
- **labour force survey:** a measure of unemployment that includes those people who are out of work without a job, want a job and either have actively sought work in the last four weeks and are available to start work in the next two weeks or have found a job and are waiting to start it in the next two weeks

Another way of expressing the number of people in a country who are available to work is to refer to the **working population**. There will always be some people who are too young to work, or who stay on in education beyond the school leaving age, or who are too old to work. All these people determine the **dependency ratio** of a country.

The participation rate

The labour **participation rate** is an indication of the people in a country who are either in work or officially registered as unemployed. It gives an indication of the extent to which the population of a country is economically active.

Difficulties involved in measuring unemployment

There is no one agreed method of measuring unemployment. This is significant because different ways of measuring unemployment can give different results. Most countries use one of the two following methods.

The claimant count

The **claimant count** is one method of measuring unemployment in a country. This is where the number of people who officially register as unemployed is counted; they register so that they are eligible to claim any benefits (known as ‘transfer payments’) that the state may provide for those people who are unemployed.

The labour force survey

An alternative way of measuring the number of people who are unemployed in a country is to identify the number of people available for work, and seeking work, but without a job. This is known as a **labour force survey**.

Comparing the claimant count and the labour force survey

» There are specific difficulties with the claimant count method of measuring unemployment. For example, it can include some people who are claiming benefit, but who are not actually available or prepared to work. Also, it excludes some people who would like to work, and who are looking for work, but who are not eligible for unemployment benefit, such as women returning to the labour force after childbirth.

» The claimant count does not meet the ILO definition of unemployment, because of the problems associated with it, and so has largely been replaced by the ILO unemployment rate, a measure based on the labour force survey. The ILO unemployment rate is a measure of the percentage of the workforce who are without jobs, but are available for work, willing to work and looking for work.
» The two methods can give different figures of unemployment; in many countries, the labour force survey measure has exceeded the claimant count measure by as much as 2% or 3% of the labour force.
» The labour force survey measure is subject to the same problems as other surveys: there is the potential for sampling error/bias in the data collection.
» No measure of unemployment is likely to be completely accurate because there will be some people out of work who are not recorded in the official statistics.

🔗 you need to demonstrate that you understand that there are two different ways to measure the number of unemployed people in an economy: the claimant count and the labour force survey. However, the claimant count does not meet the International Labour Organisation (ILO) definition of unemployment and so has been replaced by the ILO unemployment rate, a measure based on the labour force survey, as the official measure of unemployment in most countries.

The causes and types of unemployment

Unemployment in an economy can be caused by a number of factors and this gives rise to a variety of different types of unemployment:

Type of unemployment	Explanation of cause
Structural	Workers are unemployed as a result of a change of demand in an economy; this creates a change in the country's economic structure, and the declining industries will not need to employ as many people, causing a loss of jobs.
Regional	Sometimes these declining industries are concentrated in particular areas of a country and this can contribute to regional unemployment; this is especially the case where workers lack the skills and training to move from one job to another.
Cyclical (or demand deficient)	Unemployment arising from a deficiency in aggregate demand, meaning that there are insufficient jobs in the economy.
Frictional	At any one moment in an economy, some people will be between jobs (i.e. they have left one job and are waiting to start another); this type of unemployment reflects dynamic change in an economy, with some sectors expanding while others are declining. It is possible to distinguish between three different types of frictional unemployment: • search unemployment, where people are prepared to keep looking for the best possible job rather than take the first one offered • casual unemployment, where certain types of work are not regular and so at any one time some people will be out of work (e.g. in the acting profession) • seasonal unemployment, where people are out of work when it is 'out of season' (e.g. in the tourism or agriculture industry)

Technological	Technological unemployment occurs when developments in technology and working practices cause some workers to lose their jobs due to a lack of necessary skills.
Real wage (or classical)	This is where people are out of work because real wages in an economy are too high (e.g. because of the power of trade unions).
Disguised	This is where there are people who do not have a job, but who are not registered for unemployment benefits; they are not included in the claimant count of the number of unemployed people in a country.

KEY SKILL

Evaluation: you will need to be able to compare different types of unemployment and come to a judgement as to whether one type is likely to have a more damaging effect on an economy than another.

The consequences of unemployment

Unemployment has a number of consequences, including the following:

Unemployment has a number of consequences, including the following:

» Economic resources are scarce, so any unemployment is a waste of scarce resources; an economy will be underperforming and the level of national output produced will be lower than would otherwise have been the case.
» A government will lose out on potential tax revenue, in terms of both direct taxes, such as income tax, and indirect taxes, such as goods and services tax.
» A government will also find its fiscal situation worsened, not only because of the likely reduction in revenue from taxation, but also because of the money that will need to be paid out in the form of benefits to those who are unemployed. A government might also need to provide training in order to make unemployed people more employable in the future, and this will be an additional cost.
» Higher levels of unemployment can be associated with a number of social problems, such as an increase in the crime rate of a country.

🔗 Make sure that you read a question on unemployment carefully, to ensure that you do not confuse the causes of unemployment with the consequences.

4.6 Price stability

The definition of inflation, deflation and disinflation

Inflation

- **inflation:** a general increase in the average level of prices in an economy over a period of time

Inflation refers to the situation of a rise in the general level of prices in an economy over a period of time.

🔗 Candidates need to realise that it is not necessary for *all* prices in an economy to be rising to constitute a situation of inflation. It is often the case that, at any one time, the prices of some products will be falling while the prices of other products are rising. What is important in an inflationary situation is that the *general* level of prices on average is rising.

Degrees of inflation

Inflation, as has been stated, is a situation of an increase in the average price level in an economy, but there are varying degrees of inflation. In many economies, the government aims to keep the rate of inflation down to about 2%. However, in some economies a rate of inflation of 4% or 5% would not be regarded as a major problem.

In some economies, however, the rate of inflation has been significantly higher than this. In some cases, countries have experienced a rate of inflation of over 1,000%. The following are examples of countries that have experienced very high rates of inflation in the last 100 years:

» Germany in the early 1920s
» Hungary in the mid-1940s
» Zimbabwe in the mid-2000s
» Venezuela in 2017–21

Types of inflation:

Creeping inflation

- **creeping inflation:** a situation where the rate of inflation is reasonably low, say about 2%

When the rate of inflation in an economy is relatively low, say about 2%, and is reasonably stable over a period of time, the situation is one of **creeping inflation**.

Accelerating inflation

- **accelerating inflation:** a situation where the rate of inflation is rising over a period of time

Accelerating inflation is where the rate of inflation in an economy is getting significantly higher and is becoming a major problem in the economy as the real value of a currency is significantly eroded.

Hyperinflation

- **hyperinflation:** a situation where the rate of inflation is becoming very high and is damaging confidence in the country's economy

It is difficult to be precise as to exactly when an economy has reached a situation of **hyperinflation**, but it is generally where the rate of inflation has reached such a high level that it affects confidence in the economy. It may even lead to the collapse of a country's currency. Two examples of this are as follows:

» **Germany**: the rate of inflation in Germany reached over 20,000% in 1923, leading to the replacement of the currency.
» **Zimbabwe**: the rate of inflation reached over 230 million % in 2008, and this caused many people in Zimbabwe to cease using the Zimbabwe dollar and to replace it with the US dollar.

Candidates should understand that it is not possible to give a precise percentage figure to indicate when a situation of hyperinflation exists in an economy.

Deflation

- **deflation:** a general decrease in the average level of prices in an economy over a period of time

It is not inevitable that a country will always be faced with inflation. It is possible that a country could experience a situation of **deflation**, or falling average prices. This has happened in Japan in recent years.

🔗 Candidates can sometimes confuse *inflation* and *deflation* in their examination answers. It is important to distinguish clearly between a situation of inflation, which refers to rising prices, and a situation of deflation, which refers to falling prices.

Disinflation

- **disinflation:** a general increase in the average level of prices in an economy over a period of time, where the rate of increase is less than in the previous time period

» **Disinflation** refers to a situation in an economy where there is a fall in the rate of inflation. The general level of prices in the economy is still rising, but it is now rising more slowly than before.
» For example, if the rate of inflation in an economy was 5% in one year, and in the next year it fell to 3%, this would be an example of disinflation because although the general level of prices is continuing to rise, it is rising at a slower rate than in the previous year.

🔗 Candidates need to realise that in a situation of disinflation, the general level of prices in an economy is still rising. It is just that the *rate of increase* of prices is not as high as it was in the previous year.

The measurement of changes in the price level

- **general price level:** the average level of prices of all consumer goods and services in an economy at a given time
- **cost of living:** the cost of a selection of goods and services that are consumed by an average household in an economy at a given time

Changes in the **general price level** of an economy give an indication of the average price level of the various consumer goods and services at a given time. It is a way of measuring the **cost of living** in an economy.

The general price level in an economy is measured through the use of an index. There are a number of different price indices that can be used to measure changes in the cost of living in an economy over a period of time, but the most common one is the **consumer price index** (CPI).

The Consumer Price Index (CPI)

- **consumer price index (CPI):** a way of measuring changes in the prices of a number of consumer goods and services in an economy over a period of time
- **sampling:** the use of a representative sample of goods and services consumed in an economy to give an indication of changes in the cost of living
- **household expenditure:** a survey is taken on a regular basis (usually every month) to record changes in the prices of a selection of goods and services that constitutes a representative basket
- **weights:** the items in a representative sample of goods and services bought by people in an economy will not all be of the same importance; weights are given to each of the items to reflect the relative importance of the different components in the basket, and so a price index involves a weighted average
- **base year:** a year chosen so that comparisons can be made over a period of time; the base year for an index is given a value of 100

There are a number of components involved in the construction of a price index:

Components	Explanation
Basket of goods and services	A **sample** of a number of goods and services is included in a representative basket; the number of items is often around 600–700.
Household expenditure	A survey is carried out, usually every month, which monitors changes in the prices of these goods and services.
Weights	An index is constructed to show the changes in these prices, but it is recognised that some items are more important than others; these differences in importance are reflected in **weights** that are given to the different items in the basket. The more important the item, the greater the relative weight assigned in the index.
Base year	To show the changes in the general level of prices over a period of time, a starting date is chosen; this is called a **base year** and is given an index value of 100.
Calculation of the index figure	The percentage change in price for each item is then multiplied by its weight to give the average change in the index; this is usually calculated each month and shows the change in the general price level over the previous 12 months.

Possible difficulties in measurement

There are a number of possible difficulties in constructing and using a price index, such as a consumer price index, to measure changes in the general price level of an economy:

» Changes in the quality of products mean that price rises may reflect not inflation, but improvements in the quality of the products.
» One-off shocks, such as a rise in oil prices, will lead to a higher rate of inflation, but the price rise may only be temporary and so it can give a misleading measurement.
» There are various price indices that can be used to measure inflation in an economy and they are all slightly different — for example, some countries use both a consumer price index and a retail price index and there is often a difference between the two measures. So, the rate of inflation in an economy will depend partly on which price index is used to measure it.
» Different groups of people in an economy can have different inflation rates that apply to them — for example, young people could experience a different rate of inflation than old people depending on the goods and services bought by each group.
» The basket of goods and services can become outdated. Although the basket is regularly updated, this may take some time.

The distinction between money values and real data

- **nominal value:** the value of a sum of money without taking into account the effects of inflation
- **real value:** the value of a sum of money after taking into account (removing) the effects of inflation

It is important to distinguish between nominal value and real value:

» **Nominal value:** the nominal value of a given sum of money does not take into account the effects of inflation.
» **Real value:** the real value of a given sum of money shows the value after the effects of inflation have been removed.

An example of nominal value is if a person receives a wage rise of 8%. If, at that time, inflation is 7%, to obtain the real value of the wage increase, this 7% has to be subtracted from the 8% nominal increase. The real value of the wage increase is therefore 8% − 7% = 1%.

🔗 It is important that candidates can demonstrate they understand what is meant by a 'real value'.

KEY SKILL

Numerical skill: you need to be able to calculate the real value of a nominal wage increase to take into account the impact of inflation.

The causes of inflation

1. Demand-pull inflation

- **demand-pull inflation:** a rise in the general level of prices in an economy, caused by too much demand for goods and services
- **monetary inflation:** a rise in the general level of prices in an economy, caused primarily by too much money in an economy

One cause of inflation is a situation where there is too much demand in an economy — that is, aggregate demand is greater than aggregate supply. Where an increase in demand cannot be met by an increase in output, the general price level will rise — a situation of **demand-pull inflation**.

The excess demand in an economy can be due to monetary factors. An excessive increase in the money supply can make such an increase in demand possible.

Monetary inflation is closely associated with the Monetarist perspective on economics, which stresses the importance of the quantity theory of money.

2. Cost-push inflation

- **cost-push inflation:** a rise in the general level of prices in an economy, caused primarily by a significant rise in the costs of production
- **anticipated inflation:** the expected future rate of inflation in an economy
- **unanticipated inflation:** the actual rate of inflation in an economy minus the anticipated or expected rate of inflation
- **imported inflation:** a rise in the general level of prices in an economy, caused primarily by a significant increase in the price of imports

» Whereas demand-pull inflation is essentially brought about from the demand side, **cost-push inflation** is due to an increase in the costs of production that is passed on to consumers in the form of higher prices. Labour costs, for example, could rise significantly, forcing firms to raise the prices of products.
» There is also an element of **anticipated inflation** in relation to labour costs. If inflation is anticipated or expected, workers will build this into their wage demands.
» It is therefore possible that anticipated inflation actually contributes to the inflation. In such a situation, inflation becomes self-perpetuating and a self-fulfilling prophecy.
» The rate of inflation excluding anticipated inflation is known as **unanticipated inflation**.

» Cost-push inflation can sometimes be seen as **imported inflation**. This stresses the importance of an increase in the costs of imports in particular. This can either be in the form of imported raw materials and component parts that are used in domestic production or the higher prices of imported finished goods.

The consequences of inflation

- **menu costs:** the costs of continually having to change the prices of goods and services as a result of an inflationary situation
- **shoe leather costs:** in a situation of very high inflation, people need to conduct searches to ensure that their money is gaining interest; these searches are costly in time and effort
- **fiscal drag:** the idea that more people will be dragged into higher tax brackets in a situation of inflation if tax allowances are not increased in line with the rate of inflation
- **stagflation:** a situation where an economy is experiencing high inflation and high unemployment at the same time

the main negative and positive consequences of inflation for an economy:

Negative consequences	Positive consequences
• The purchasing power of a given sum of money will fall (i.e. it will be able to buy less in real terms). • A country's exports will become uncompetitive, although the effect of this will depend on the rate of inflation in one country compared to that in another. • There will be a redistribution of income, with some people adversely affected, such as those on fixed incomes, and creditors (people who lend money) unless the loans are index linked to the rate of inflation. Savers will also be negatively affected, unless they can find a savings account with a rate of interest that is greater than the rate of inflation. • **Menu costs** — repeatedly changing the prices that are being advertised for various products can be expensive and time consuming. • **Shoe leather costs** — searching continually for the best possible returns to try to keep ahead of inflation also involves extra time and effort. • Uncertainty in investment — some firms may be reluctant to plan investment as inflation creates a great deal of uncertainty in an economy.	• A relatively low rate of inflation, caused by an increase in aggregate demand, could lead to people feeling more optimistic about future economic prospects. • If prices rise by more than costs, this will help to increase the profits of firms, assuming that demand is price elastic. • Redistribution of income may work in favour of some people in an economy, such as borrowers; inflation will make their debt less in real terms. This may encourage people to spend more, improving their standard of living. • If the real cost of debt is reduced, but the price of property rises in line with inflation, property owners will benefit significantly. • If firms experience an increase in their profits, they may be encouraged to expand and this could lead to a reduction in unemployment (although high inflation can also be associated with high unemployment, a situation known as **stagflation**).

• A situation of **fiscal drag** can occur if tax allowances are left unchanged in an inflationary situation; fiscal drag is where inflation moves taxpayers into higher tax brackets.	

It should be clear, therefore, that the overall effects of inflation will depend on a number of factors. These include:

» Whether the inflation is anticipated or unanticipated
» The extent of the increase in prices
» Whether the inflation rate is accelerating or creeping
» The extent to which other countries are experiencing inflation

🔗 Candidates often make broad generalisations in examinations in relation to the possible effects of inflation. You need to make sure that you demonstrate an understanding that the effects of inflation in any particular country will depend on a number of possible factors.

KEY SKILL

Evaluation: you will need to be able to come to a judgement in terms of an overall assessment of the consequences of inflation, recognising that these can be both positive and negative. It is important you realise that there are potential benefits and costs of inflation.

5 Government macroeconomic intervention

5.1 Government macroeconomic policy objectives

There are a large number of different government macroeconomic policy objectives, but this chapter focuses on the use of government policy (i.e. fiscal policy, monetary policy and supply-side policy) to achieve just three of these macroeconomic objectives:

» price stability
» low unemployment
» economic growth

🔗 Knowledge of policy conflicts and trade-offs will not be required at AS Level.

5.2 Fiscal policy

The meaning of a government budget

- **fiscal policy:** the use of public revenue and/ or public expenditure to influence the level of aggregate demand in an economy
- **government budget:** a financial statement that outlines a government's proposed revenue and expenditure for a specific period of time

Fiscal policy is the deliberate use of revenue and expenditure decisions by a government to influence economic activity in a country, especially the level of aggregate demand. The idea is that by making changes to decisions about revenue and expenditure, an economy can be 'fine-tuned' to achieve particular aims.

A **government budget** is a financial statement that outlines a government's proposed revenue/income and expenditure/spending for a specific period of time, usually the next financial year.

🔗 **Time:** a budget is a financial statement that covers a particular period of time, usually 1 year. You may need to refer to the time period when discussing government policies, such as when analysing fiscal policy in relation to government expenditure and tax revenue. Time is also important in medium- and long-term financial planning, such as in terms of how long it may take for fiscal policy, monetary policy or supply-side policy to take effect.

The distinction between a government budget deficit and a government budget surplus

In each economy in the world, the government will produce its accounts in the form of a summary of its income and expenditure.

A government can deliberately plan for one of three financial positions to achieve its objectives:
» a budget deficit
» a budget surplus
» a balanced budget

Budget deficit

- **budget deficit:** where projected government revenue is less than planned government expenditure

» In the case of a **budget deficit**, the projected revenue is less than the planned expenditure (i.e. more money is spent than is received). In such a situation, the government is injecting more into an economy than it is withdrawing.
» For example, if an economy needs to be stimulated in order to reduce unemployment and to increase economic growth, a government can plan for a budget deficit whereby taxation is reduced and government expenditure increased.

Budget surplus

- **budget surplus:** where projected government revenue is greater than planned government expenditure

» In the case of a **budget surplus**, the projected revenue is greater than the planned expenditure (i.e. less money is spent than is received). In such a situation, the government is taking more out of an economy than it is putting in.
» For example, if an economy needs to be deflated in order to reduce inflation, a government can plan for a budget surplus whereby taxation is increased and government expenditure reduced.

Balanced budget

- **balanced budget:** where projected government revenue and planned government expenditure are equal

In the case of a **balanced budget**, the projected revenue, such as from taxation, is exactly equal to the government's planned expenditure. In this sense, where both sides of the budget are equal, it is neutral.

KEY SKILL

Evaluation: you need to be able to discuss why a government might deliberately plan for a budget deficit or a budget surplus rather than for a balanced budget. For example, a government might deliberately plan for a budget deficit in an expansionary fiscal or monetary policy to reduce unemployment in an economy. Alternatively, a government might deliberately plan for a budget surplus in a contractionary fiscal or monetary policy to reduce inflation in an economy.

The meaning and significance of the national debt

- **national debt:** the total of all debt accumulated by a government

The **national debt** refers to the total of all debt that has been accumulated over a period of time by the government or the public sector of a country — in other words, it is a government's stock of outstanding debt.

Changes in the size of a country's national debt can be brought about by government policy:

» **Increase in the size of the national debt:** if there is a relatively high rate of unemployment in an economy, a government may decide to plan for a budget deficit and the effect of this will be to increase the size of the national debt.
» **Decrease in the size of the national debt:** if there is a high rate of inflation in an economy, a government may decide to plan for a budget surplus and the effect of this will be to reduce the size of the national debt.

🔗 It is important to understand that at a time when the size of a budget deficit is falling, the size of the national debt will still be rising because as long as there is a budget deficit, a government will be spending more than it is receiving. It is important also to understand that a budget surplus does not necessarily mean a reduced national debt. A government would need to choose to use the budget surplus for the purpose of reducing the size of the national debt.

🔗 It is important that the national debt of a country is distinguished from debt that is accumulated in one financial year. The national debt refers to the debt that has been accumulated over many years.

🔗 It is important that candidates are able to distinguish between a stock of money and a flow of money. A country's national debt is a stock of money that has been accumulated over many years. A budget deficit refers to a flow of money in one financial year. If a government is successful in reducing the size of a country's budget deficit in a particular year, the size of the country's national debt may still increase during the year.

🔗 **Time:** a country's national debt, like its government budget, is linked to the concept of time. Whereas a budget usually covers a relatively short period of time, such as a year, a country's national debt will have accumulated over a much longer period of time, involving many years and even centuries in some cases.

KEY SKILL

Analysis: you need to be able to analyse a government budget and a national debt in terms of a distinction between a flow of money and a stock of money.

Taxation

Different types of tax

- **direct tax:** a tax that is imposed on the incomes of individuals and firms; examples are income tax (on the incomes of individuals), corporation tax (on the profits of companies) and inheritance tax (on the wealth of individuals)
- **income tax:** a direct tax on the incomes of individuals
- **indirect tax:** a tax that is imposed on expenditure; it is indirect in that the tax is only paid when the product on which the tax is levied is purchased
- **specific tax:** an indirect tax that is a fixed amount per unit of output
- **ad valorem tax:** an indirect tax with a percentage rate (e.g. a tax rate of 20% per product sold)

It is possible to distinguish between different types of tax: some will be direct, while others will be indirect, and they can also be distinguished as to whether they are progressive, regressive or proportional.

Direct and indirect taxes

A **direct tax** is a tax imposed on the incomes of individuals and the profits of firms.
Examples of direct taxation include the following:

» **Income tax** (on the incomes of individuals). There is usually a personal allowance, which is tax free, and then different tax rates for different levels of income over the tax-free allowance. Income tax rates typically increase as incomes rise.
» **Corporation tax**, or corporate tax (on the profits of firms). In some countries, the tax rate will vary depending on the size of a firm's profits, while in other countries, the tax rate will be the same for all firms.
» **Inheritance tax:** a direct tax on those people who have inherited money, property or possessions from someone who has died. There is usually a threshold at which the tax is payable depending on the value of the estate that has been inherited.

An **indirect tax**, on the other hand, is a tax imposed on expenditure. Indirect taxes can take different forms:

» **Specific tax:** a specific tax on a product, such as an excise duty, involves a requirement to pay a specific amount per unit in tax.
» **Ad valorem tax:** an ad valorem tax, on the other hand, requires a percentage of the selling price to be paid in tax. For example, a tax such as value added tax (VAT) or goods and services tax (GST) will require a particular percentage to be paid in tax, such as 20%.

It is important to distinguish between the incidence and the impact of a tax:

» **Incidence of tax:** this refers to how the burden of taxation is distributed between producers and consumers. For example, with an indirect tax on expenditure, the burden of the tax is likely to be shared between the producer and the consumer. However, if the price elasticity of demand is perfectly inelastic, and there is a vertical demand curve, the incidence of the tax will be entirely on the consumer. On the other hand, if the price elasticity of demand is perfectly elastic, and there is a horizontal demand curve, the incidence of the tax will be entirely on the producer.
» **Impact of tax:** this refers to the impact on the individual or firm on which a tax is levied.

KEY SKILL

Application: you need to be able to give appropriate examples of different kinds of tax. For example, income tax is an example of a direct tax, whereas VAT is an example of an indirect tax.

Progressive, regressive and proportional taxes

- **progressive taxation:** where taxation takes a higher proportion of a person's income as that income rises
- **regressive taxation:** where taxation takes a larger proportion of low incomes than it does of high incomes
- **proportional taxation:** where taxation takes an equal proportion of income whatever a person's income level
- **flat-rate tax:** a tax with a constant marginal rate

Progressive taxation

Many governments make use of progressive taxation to achieve the macroeconomic objective of a more equal distribution of income. An income tax, for example, is usually a progressive tax as the tax rate rises as the level of income rises — it not only takes more from a person as their income rises, but a higher *proportion* of that income.

Regressive taxation

Whereas a progressive direct tax, such as income tax, has different rates of tax depending on a person's income (although, in theory, direct taxation can also be regressive), a regressive indirect tax, such as a goods and services tax, is paid at a constant rate. In such a situation, all people who buy a particular good or service will pay the same percentage, so this will have a greater impact on relatively low-income as compared to relatively high-income individuals.

Proportional taxation

Another possible type of taxation is proportional taxation. This is where a tax takes an equal proportion of income from a person whatever that person's income. In this situation, the tax has neither a progressive nor a regressive effect.

An example of a proportional tax is a **flat-rate tax**, which has a constant marginal tax rate.

Marginal and average rates of taxation

- **marginal tax rate:** the proportion of an increase in income that is paid in tax
- **average tax rate:** the average percentage of total income that is paid in tax

The **marginal tax rate** refers to the proportion of an increase in income that is paid in tax. This can also be referred to as the *marginal propensity to pay tax*.

The **average tax rate** refers to the average percentage of total income that is paid in taxes. This can also be referred to as the *average propensity to pay tax*.

🔗 **The margin and decision making:** a progressive tax is one where the proportion of income taken in tax (the tax rate) rises with income. This is achieved by a system of rising marginal rates of tax with increases in income (e.g. 20%, 30%, 40% and 50%).

🔗 **The role of government and the issues of equality and equity:** progressive taxation can be used by a government deliberately to redistribute income and wealth so that the distribution becomes more equal.

Canons of taxation

- **canons of taxation:** the main principles of taxation, to which any system of taxes should adhere in order to be effective

principles or canons of taxation:

Principle	Explanation
Equity/fairness	The burden of taxation should take into account the ability to pay the tax.
Certainty/transparency	Information about taxation needs to be made available so that it is seen as transparent.
Convenience	The payment and collection of taxes need to be as convenient as possible.
Cost	The cost of administering and collecting taxes should be as low as possible.
Efficiency	Taxation should not lead to any disincentives, such as discouraging people from working.

Reasons for taxation

There are a number of reasons for taxation:

» **To provide revenue for a government:** taxation is used to enable a government to finance its expenditure on a variety of different goods and services, such as the provision of public goods (e.g. street lighting and national defence).
» **To achieve government macroeconomic policy objectives:** taxation is used to influence the level of aggregate demand in an economy — for example, if there is a high rate of unemployment, taxes could be lowered to stimulate the level of demand in an economy in an attempt to reduce the level of unemployment.
» **To redistribute income and wealth:** taxation is used to finance the provision of a range of transfer payments to those people in society who are relatively less well off.
» **To avoid negative externalities:** taxation is used to discourage firms from causing any negative impact on third parties (e.g. taxes on firms can be used to reduce the level of pollution in an economy).
» **To discourage consumption/production of demerit goods:** taxation is used in an attempt to lower the level of demand for demerit goods (e.g. tobacco and alcohol).

Government spending

- **capital (investment) spending:** government spending on fixed assets
- **current spending:** government spending on day-to-day running costs

It is important to distinguish between two types of government spending:

» **Capital (investment) spending:** this refers to government expenditure on fixed assets, such as expenditure on building a new road or on extra defence equipment.
» **Current spending:** this refers to government expenditure on the day-to-day running costs of a government, such as expenditure on the wages or salaries of public sector workers.

Reasons for government spending

» **It is a key component of fiscal policy:** government spending is one of the components of aggregate demand, so an increase in government spending (assuming the other components stay constant) will increase aggregate demand in an economy and so aid the management of the economy.

» **It is important in providing public goods:** governments spend on public goods (e.g. street lighting and national defence) because these goods would not be provided in the private sector and so governments are required to provide them.
» **It is important in providing merit goods:** governments spend on merit goods (e.g. education and healthcare) because, although these goods could be provided in the private sector, they are likely to be under-consumed and so governments spend money on them to encourage their consumption.
» **It helps to achieve greater equity in an economy:** a government could decide to spend money on a range of benefits and transfer payments as a way of achieving greater equity (e.g. state pensions and unemployment benefits).

🔗 **Progress and development:** government spending can contribute to the progress and development of an economy, such as through spending on public goods (e.g. defence and police) and on merit goods (e.g. education and healthcare).

🔗 **The role of government and the issues of equality and equity:** one reason for government spending is that it could help to achieve greater equality and equity in an economy (e.g. through spending on a range of transfer payments).

Discretionary fiscal policy and automatic stabilisers

- **discretionary fiscal policy:** the use of deliberate changes in taxation and/or public expenditure with the intention of changing the level of aggregate demand in an economy
- **automatic stabilisers:** where changes in the level of taxation and/ or public expenditure automatically bring about changes in an economy without the need for deliberate action by a government

It is important to distinguish between **discretionary fiscal policy** and **automatic stabilisers**.

Discretionary fiscal policy

Sometimes a government will deliberately change taxation and/or public expenditure to bring about a desired change in the level of economic activity in the economy. This is known as discretionary fiscal policy.

Automatic stabilisers

» Whereas discretionary fiscal policy involves deliberate action by a government, automatic stabilisers refer to a situation where changes in an economy take place without the need for deliberate government action.
» For example, in a recession, when the level of unemployment in an economy is likely to rise, the revenue received by a government from taxation is likely to fall, as fewer people are working, while at the same time government expenditure is likely to rise, such as in the form of unemployment benefits to those out of work. This results in ‘automatic stabilisation’ because the increased government spending will impact on the trend rate of economic growth.
» Automatic stabilisers will be discussed in Chapter 9, section 9.2, in relation to the business (trade) cycle.

The distinction between expansionary and contractionary fiscal policy

- **expansionary fiscal policy:** one that causes aggregate demand in an economy to increase
- **contractionary fiscal policy:** one that causes aggregate demand in an economy to decrease

As previously explained, fiscal policy is the use of taxation and/or public expenditure to influence the level of aggregate demand in an economy. It is important to distinguish between expansionary and contractionary fiscal policy:

» **Expansionary fiscal policy** is where a government decides to increase its expenditure and/or reduce taxation to boost the level of aggregate demand in an economy. This approach will be appropriate if the objective is to encourage economic growth, lower the rate of unemployment or increase the rate of inflation.
» **Contractionary fiscal policy** is where a government decides to lower its expenditure and/or increase taxation to reduce the level of aggregate demand in an economy. This approach will be appropriate if the objective is to lower economic growth or the rate of inflation.

AD/AS analysis of the impact of expansionary and contractionary fiscal policy

AD/AS analysis can be used to assess the impact of expansionary fiscal policy and contractionary fiscal policy on equilibrium national income in relation to:

» the level of real output
» the price level
» the employment level

The impact of expansionary fiscal policy

» In Chapter 4, section 4.3, it was stated that if there is a change in any of the components of aggregate demand, such as a change in consumption expenditure (*C*), investment expenditure (*I*), government expenditure (*G*) or net expenditure on exports (*X* − *M*), the *AD* curve will shift. If there is an increase in aggregate demand, the curve will shift to the right.
» This can be seen in Figure 4.3 on page 67, where there is a shift in the aggregate demand curve from *AD*0 to *AD*1. This results in an increase in the level of real output and employment (from *Y*0 to *Y*1) and an increase in the price level (from *P*0 to *P*1).
» An expansionary fiscal policy could bring this about in a number of ways. For example, if a government decides to raise its expenditure, this will increase *G*. If it decides to lower income tax, this could increase *C*.

The impact of contractionary fiscal policy

» If there is a decrease in aggregate demand, the *AD* curve will shift to the left. This will cause a decrease in the level of real output, employment and the price level.
» A contractionary fiscal policy could bring this about in a number of ways. For example, if a government decides to lower its expenditure, this will reduce *G*. If it decides to increase income tax, this could reduce *C*.

🔗 **Equilibrium and disequilibrium:** the position of equilibrium and disequilibrium, as shown through the interaction of *AD* and *AS*, can be affected by expansionary or contractionary fiscal policy.

5.3 Monetary policy

The definition of monetary policy

- **monetary policy:** the use of interest rates and/or the money supply to influence the level of aggregate demand in an economy

Monetary policy is concerned with how the price and/or the quantity of money can be used to influence the level of aggregate demand in an economy. It refers to actions that a country's

government, central bank or monetary authority can take to influence how much money and credit is in an economy through the manipulation of the rate of interest and how much it costs to borrow that money.

The tools of monetary policy

- **hot money:** flows of money that move from one country to another to take advantage of higher rates of return in various countries
- **quantitative easing:** the process whereby the government, central bank or monetary authority of a country deliberately buys bonds and bills in order to increase the money supply in an economy
- **open market operations:** the process of a government, central bank or monetary authority buying or selling bonds in order to influence the money supply in an economy

There are three main tools of monetary policy:

» interest rates
» money supply
» credit regulations

1. Interest rates

» The *price* of money refers to the interest rate.
» For example, a government, central bank or monetary authority could reduce the interest rate in an economy to stimulate employment, but this could lead to an increase in the rate of inflation in that economy.
» If a government decides to increase the interest rate in an economy, this will attract '**hot money**' into the country, raising the value of the exchange rate and making export prices more expensive and import prices less expensive, leading to the possibility of an adverse effect on the balance of payments.

2. Money supply

» A second monetary approach is to influence the quantity of money in an economy (i.e. the money supply).
» One way of doing this is through changes in credit regulations — for example, if credit regulations are tightened, the money supply is likely to be reduced.
» Many countries, since the financial crisis of 2007–08, have decided to increase the money supply in the economy through a process of **quantitative easing**. This is where the government buys bonds and bills (i.e. securities), giving financial institutions more liquidity and so increasing the money supply in the economy. This process of a government buying bonds is known as **open market operations**.

3. Credit regulations

» Credit regulations refer to laws that relate to borrowing money on credit. They are concerned with loans and/or hire purchase agreements.
» The regulations cover the information consumers should be provided with before they enter into a credit agreement, the content and form of credit agreements, the method of calculating the rate of interest and the procedures relating to default, termination and early settlement.
» As has already been pointed out, changes in credit regulations can impact the money supply. If credit regulations are loosened, the money supply is likely to be increased.

The distinction between expansionary and contractionary monetary policy

- **expansionary monetary policy:** one that causes aggregate demand in an economy to increase
- **contractionary monetary policy:** one that causes aggregate demand in an economy to decrease

As previously explained, monetary policy is the use of interest rates and/or the money supply to influence the level of aggregate demand in an economy. It is important to distinguish between expansionary and contractionary monetary policy.

» **Expansionary monetary policy** is where a government decides to increase the money supply and/or lower interest rates to boost the level of aggregate demand in an economy. This approach will be appropriate if the objective is to encourage economic growth or to lower the rate of unemployment.
» **Contractionary monetary policy** is where a government decides to decrease the money supply and/or increase interest rates to reduce the level of aggregate demand in an economy. This approach will be appropriate if the objective is to lower economic growth or the rate of inflation.

AD/*AS* analysis of the impact of expansionary and contractionary monetary policy

AD/*AS* analysis can be used to assess the impact of expansionary monetary policy and contractionary monetary policy on equilibrium national income in relation to:

» the level of real output
» the price level
» the employment level

The impact of expansionary monetary policy

» In Chapter 4, section 4.3, it was pointed out that if there is a change in any of the components of aggregate demand, such as a change in consumption expenditure (*C*), investment expenditure (*I*), government expenditure (*G*) or net expenditure on exports (*X* −*M*), the *AD* curve will shift.
» This can be seen in Figure 4.3 on page 67, where there is a shift in the aggregate demand curve from *AD*0 to *AD*1. This results in an increase in the level of real output and employment (from *Y*0 to *Y*1) and an increase in the price level (from *P*0 to *P*1).
» An expansionary monetary policy could bring this about in a number of ways. For example, if a government decides to increase the money supply and/or reduce the interest rate, this will increase *C*. A reduction in the interest rate is also likely to increase investment (*I*).

The impact of contractionary monetary policy

» If there is a decrease in aggregate demand, the *AD* curve will shift to the left. This will cause a decrease in the level of employment, real output and the price level.
» A contractionary monetary policy could bring this about in a number of ways. For example, if a government decides to lower the money supply and/or increase the interest rate, this will decrease *C*. An increase in the interest rate is also likely to decrease investment (*I*).

🔗 **Equilibrium and disequilibrium:** the position of equilibrium and disequilibrium, as shown through the interaction of *AD* and *AS*, can be affected by expansionary or contractionary monetary policy.

5.4 Supply-side policy

The meaning of supply-side policy

- **supply-side policy:** a policy designed to enable markets to work more efficiently

Supply-side policy is a general term referring to a number of different actions that a government can take to improve the efficiency of markets in an economy. It is different in approach from the other two types of macroeconomic policy because it approaches economic issues from the supply, rather than from the demand, side.

In particular, supply-side policy is designed to have an effect on long-run aggregate supply which can be represented by the long-run aggregate supply curve (*LRAS*):

» Supply-side policies are aimed at making markets and industries operate more efficiently so that they contribute to a faster rate of growth of real national output.
» Successful supply-side policies will have the effect of shifting the *LRAS* curve to the right, leading to a rise in the productive potential output of an economy.
» The significant advantage of an improved supply-side performance in an economy is that sustained economic growth can be achieved without causing a rise in inflation.

The objectives of supply-side policy

- **productivity:** the measurement of output per unit of input per period of time
- **productive capacity:** the maximum potential output of an economy
- **labour productivity:** productivity measures the level of efficiency in the use of resources; labour productivity, therefore, measures the efficiency of labour in terms of the output per person per period of time

The objectives of supply-side policy can include the following:

» Increasing **productivity**: supply-side policies, with their emphasis on bringing about greater efficiency in the production process, can contribute to a greater rate of output (i.e. products produced) per unit of input (e.g. labour and capital).
» Increasing **productive capacity**: supply-side policies can contribute to an increase in the maximum possible output of an economy.

Labour productivity is the measurement of the efficiency of labour in terms of the output per worker per period of time.
The productivity of workers can vary for several reasons, including differences in:

» education
» training
» skills
» experience
» technical knowledge
» availability of capital
» working methods and practices
» motivation

KEY SKILL

Analysis: you need to be able to analyse how labour productivity can be increased in an economy over a period of time. For example, improvements in education and training will lead to a workforce that is better qualified and appropriately skilled, resulting in an increase in the maximum possible output of an economy.

🔗 **Time:** you need to understand that increases in productivity and productive capacity in an economy will occur in the long run as a result of implementing different supply-side policies.

🔗 It is important to distinguish between productivity and production. Candidates often confuse these two terms. 'Production' refers to total output from resources, whereas 'productivity' refers to the efficiency of an input (e.g. labour) into the production process.

The tools of supply-side policy

There are two main types of supply-side policy:

» Market-based supply-side policies to increase competition and efficiency
» Interventionist supply-side policies to overcome market failure

Market-based supply-side policies	Interventionist supply-side policies
• Increasing incentives to work by lowering income tax and unemployment benefits • Reforming trade unions so that labour is not in such a powerful bargaining position with employers; this could contribute to a reduction in the number of days lost through industrial action in an economy • Encouraging privatisation in an economy so that firms, needing to make a profit to survive, become more efficient • Encouraging deregulation, such as through a reduction in the barriers to entry into an industry, which will allow more private firms to enter a market • Encouraging an increased level of competition in markets, such as through tax incentives	• Increasing expenditure on education to improve the quality, and therefore the productivity, of the labour force • Increasing expenditure on training to enable workers to learn new skills and move more easily from one type of work to another if they become unemployed; this is likely to improve the flexibility and occupational mobility of labour in an economy • Providing more information about job vacancies in different parts of a country; this is likely to improve the geographical mobility of labour • Encouraging infrastructure development, such as the construction of new road and rail links to improve transport and reduce costs • Support for technological improvement, such as providing financial incentives for research and development

AD/*AS* analysis of the impact of supply-side policy

» *AD*/*AS* analysis can be used to assess the impact of supply-side policy on equilibrium national income in relation to the level of real output, the price level and the employment level.
» In Chapter 4, section 4.3, it was pointed out that equilibrium national income can be affected by changes in the *AS* curve as well as the *AD* curve. If there is an increase in aggregate supply, as a result of supply-side policies, the *AS* curve will shift to the right. This can be seen in Figure 4.5 on page 68, where the curve shifts from *AS*0 to *AS*1. If there is a decrease in aggregate supply, the curve will shift to the left.
» In the long run, the shape of the *AS* curve will change. At low levels of output, the *LRAS* curve can be horizontal, indicating that it is perfectly elastic. At high levels of output, the *LRAS* curve can be vertical, indicating that it is perfectly inelastic (see Figure 4.8 on page 70).
» Indeed, some economists argue that the *LRAS* curve is perfectly inelastic, indicating that an economy will always operate at full capacity. That is why it is important that supply-side policy

shifts the *LRAS* curve to the right, allowing for an increase in the productive capacity of an economy.
» A shift to the right of the *LRAS* curve will increase the level of real output and the level of employment, but will decrease the price level.
» A shift to the left of the *LRAS* curve will decrease the level of real output and the level of employment, and increase the price level.

KEY SKILL

Application: you need to be able to apply supply-side policy to particular methods or tools, including examples of both market-based and interventionist supply-side policies.

🔗 **Equilibrium and disequilibrium:** the position of equilibrium and disequilibrium, as shown through the interaction of *AD* and *AS*, can be affected by supply-side policy. For example, supply-side policies will be able to shift the *LRAS* curve to the right, leading to an equilibrium position with a higher level of real output and employment and with a lower price level.

6 International economic issues

6.1 The reasons for international trade

The distinction between absolute and comparative advantage

- **absolute advantage:** a situation where a country can produce a particular good or service using fewer resources than another country
- **comparative advantage:** a situation where a country can produce a good or service relatively more efficiently (at a lower domestic opportunity cost) than another country

International trade between countries is based on specialisation, whereby one country is more efficient than another at producing a particular product. This gives rise to two types of advantage:

» **Absolute advantage**
» **Comparative advantage**

Absolute advantage

» Absolute advantage refers to a situation where one country is able to produce a particular good with fewer resources than another country.
» As a result of this, the country will enjoy a cost advantage. This is why absolute advantage is often referred to as ‘absolute cost advantage’.
» The reason for this advantage is that each country is endowed with a particular mix of factors of production, so one country may be more efficient than another country at producing a particular good.

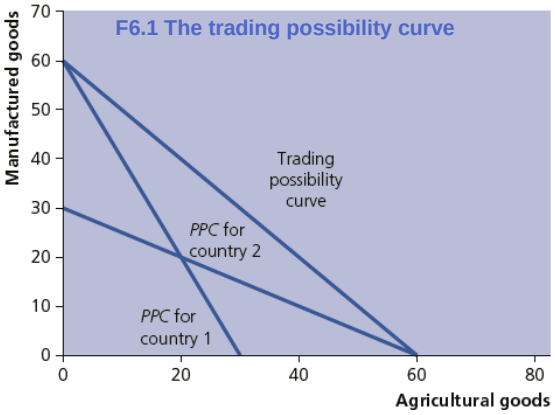
KEY SKILL

Application: a country in the Caribbean, such as Jamaica, has an absolute advantage over the UK in the production of bananas because the weather conditions are much more favourable. This makes Jamaica more efficient at producing bananas than the UK, and so gives Jamaica a cost advantage.

Comparative advantage

- **trading possibility curve**: a means of showing the potential advantages of two countries trading with each other, as long as the opportunity costs of production are different

» Comparative advantage takes into account not just the absolute efficiency of one country compared to another, but its relative efficiency.
» This then allows a country to produce something in which it has a lower opportunity cost than another country. This can be seen in Figure 6.1:
» Two countries, country 1 and country 2, have different comparative advantages, which can be seen by the slope of their respective production possibility curves (*PPCs*).
» Country 1 has a comparative advantage in the production of manufactured goods whereas country 2 has a comparative advantage in the production of agricultural goods.



» In the absence of international trade, each country is constrained to consume along its *PPC*.
» If, however, country 1 specialises in the production of manufactured goods and country 2 specialises in the production of agricultural goods, and if trade takes place on a one-to-one basis (i.e. one unit of manufactured goods is exchanged for one unit of agricultural goods), this expands the consumption possibilities for both countries.
» As a result of applying the principle of comparative advantage, the **trading possibility curve** shows the potential consumption points for each country in this situation. This curve, shown in Figure 6.1, illustrates the importance of opportunity cost in international trade.
» The principle of comparative advantage is reflected in differences in opportunity cost. World trade will increase substantially if each country concentrates on producing the product(s) in which it has the lowest opportunity cost.

Figure 6.1 shows the situation where just two countries are trading with each other. This is known as **bilateral trade**. in the real world, the situation is more complex than this, given the existence of **multilateral trade**. Multilateral trade has been made more likely by the process of **globalisation**. This has been made possible by a number of factors, including progress in trade liberalisation.

- **bilateral trade:** where trade takes place between two countries
- **multilateral trade:** a more realistic situation than bilateral trade, where trade takes place between a number of countries
- **globalisation:** the process whereby there is an increasing world market in goods and services, making an increase in multilateral trade more likely; it has been made possible by a number of factors, including progress in trade liberalisation

🔗 It is important that candidates can demonstrate their understanding of the distinction between *bilateral trade* and *multilateral trade*. Economists use two countries on production possibility curve diagrams, but in reality, trade will involve the participation of many more countries in the world.

The benefits of specialisation and free trade

- **specialisation:** the process whereby individuals, firms and economies concentrate on producing those products in which they have an advantage
- **free trade:** trade that is not restricted or limited by different types of import or export control
- **trade liberalisation:** the removal or reduction of restrictions or barriers to the free exchange of goods and services between countries
- **trade creation:** the creation of new trade as a result of the reduction or elimination of trade barriers

Some economists stress the potential advantages of specialisation and **free trade** for a country. These potential advantages can be achieved through a process known as **trade liberalisation**.

The advantages of specialisation and free trade include the following:

» World output can be increased.
» Resources are allocated more efficiently as a result of the greater specialisation.
» The higher output can be produced at lower average cost, which could lead to lower prices for consumers.
» Consumers can have a wider range of products to choose from.
» There can be a substantial increase in economic growth.
» By specialisation in the production of goods in which countries have a comparative advantage and lower opportunity cost, there can be an increase in economic welfare for all countries.
» It can lead to an improved standard of living.

A reduction in trade barriers will bring about a greater degree of free trade and this process of trade liberalisation will lead to **trade creation**.

🔗 **Progress and development:** international trade can lead to progress and development because the increase in world output can give consumers in different countries a wider range of products to choose from, possibly at lower prices, leading to an improvement in standards of living.

KEY SKILL

Evaluation: you will need to be able to evaluate the potential benefits of specialisation and free trade by comparing these with the arguments for protectionism. Arguments for and against protectionism are covered in section 6.2.

The World Trade Organization

- **World Trade Organization:** an organisation set up in 1995 to promote global free trade in the world through the reduction of trade barriers

The **World Trade Organization** (WTO) was set up in 1995:

» It is the only global organisation dealing with the rules of trade between countries.
» It replaced another organisation that had existed since 1948 with the same purpose of encouraging free trade, the General Agreement on Tariffs and Trade (GATT).
» The head office of the WTO is in Geneva, Switzerland, and it currently has 164 member countries.
» It has sought to regulate trade between countries through a series of discussions; the most recent round of discussions is the Doha Round and these have been taking place since 2000.
» WTO agreements are negotiated and signed by the majority of the world's trading countries and the objective of these agreements is to ensure that international trade flows as smoothly, predictably and freely as possible.
» If there are any trade disputes between countries, the WTO can help to resolve them.

KEY SKILL

Analysis: you should be able to analyse the contribution of the World Trade Organization to the growth of free trade and trade liberalisation through its role in promoting and supporting free trade, the removal of trade barriers and the establishment of multilateral trade agreements.

Exports, imports and the terms of trade

Exports are the goods and/or services that are produced domestically in one country and sold to other countries.
Imports are the goods and/or services that are produced in foreign countries and consumed by people in the domestic economy.

The measurement of the terms of trade

- **terms of trade:** the price of a country's exports in relation to the price of the country's imports

The **terms of trade** refer to the relative value of export prices and import prices. They show the amount of imports an economy can purchase per unit of exports. They are important because they show the gain to a country from taking part in international trade. They are calculated by:

$$\frac{index\ number\ showing\ the\ average\ price\ of\ exports}{index\ number\ showing\ the\ average\ price\ of\ imports} \times 100$$

The causes of changes in the terms of trade

If, in one year, the terms of trade are equal to 100 and import prices rise by more than export prices, the terms of trade will fall below 100. This is an unfavourable change because more exports will need to be sold to buy the same number of imports.

However, if in another year, the terms of trade are equal to 100 and import prices rise by less than export prices, the terms of trade rise above 100. This is a favourable change because fewer exports will need to be sold to buy the same number of imports, assuming a situation of ceteris paribus.

The impact of changes in the terms of trade

Changes in the terms of trade can have a significant impact on an economy:

» When the terms of trade fall below 100, this is regarded as an unfavourable change because more exports will need to be sold to buy the same number of imports. However, this means that exports have become relatively cheaper than imports, and if the price elasticity of demand for both exports and imports is elastic, this will help to improve a country's balance of trade situation.
» When the terms of trade rise above 100, this is regarded as a favourable change because fewer exports will need to be sold to buy the same number of imports. However, this means that exports have become relatively more expensive than imports, and if the price elasticity of demand for both exports and imports is elastic, this will be likely to worsen a country's balance of trade situation.

🔗 Candidates often confuse the terms of trade with the balance of trade. The terms of trade simply indicate the relationship between changes in the prices of exports and imports; they give no indication of the changes in the quantity or value of exports and imports that are traded between countries.

🔗 Candidates sometimes regard a fall in the terms of trade as something that is unhelpful to a country's trading position. In fact, a fall in the terms of trade means that export prices have become relatively cheaper, and if the price elasticity of demand for both exports and imports is elastic, this will help to improve a country's balance of trade situation.

KEY SKILL

Numerical skills: if the index number showing the average price of exports is 104.2 and the index number showing the average price of imports is 97.4, then the terms of trade will be (104.2/97.4) × 100 = 106.98, rounded up to 107.

The limitations of the theories of absolute and comparative advantage

The theories of absolute and comparative advantage are helpful in explaining how total world output can be increased as a result of specialisation, but there are a number of real-world limitations to the predictions of these theories. They include the following:

» It is assumed that there are no transport costs involved in international trade, but this is unrealistic — transport costs may actually offset any cost advantages arising from applying the theories.
» It is assumed that there are constant returns to scale and constant costs of production, but it is always possible that an increase in output will lead to diseconomies of scale and a situation of rising average costs of production. (Returns to scale are discussed in Chapter 7, section 7.5.)
» It is assumed that there is free trade between countries, but in reality, there are many trade restrictions that exist in different parts of the world.
» It is assumed that exchange rates are stable and that the benefits of international trade will not be affected by fluctuations in exchange rates, but movements in exchange rates can help to make trade more or less advantageous to a country in the real world.
» It is assumed that factor inputs can switch between different products easily and work with the same efficiency, but this may not necessarily happen.

Other explanations and determinants of trade flows

Traditionally, economists have focused on the theories of absolute and comparative advantage as the main explanation of international trade flows. However, given the existence of the real-world limitations already discussed, economists have begun to look at other possible explanations and determinants of trade flows. These include the following:

» **Competitive advantage:** this is the idea that it is not so much opportunity cost that is important, but the actual cost of production in different countries, such as the decision of multinational companies to locate production in particular countries to take advantage of differences in labour costs.
» **Factor endowment:** differences in endowments of factors of production in various countries are an element in the two theories, but some economists have stressed the significance of differences in the quality and quantity of factors of production between countries.
» **Government policy:** a government may be concerned about the possible disadvantages of overspecialisation in particular products and so might decide to encourage a greater degree of diversification of production than might otherwise have been the case.

6.2 Protectionism

The meaning of protectionism in the context of international trade

- **protectionism:** the restriction of free trade between countries in an attempt to protect local firms and industries from competition

Protectionism in the context of international trade refers to those policies that protect domestic producers from international competition, or give support to them. It is also used to maintain the independence and self-sufficiency of strategic domestic industries, such as energy or defence. The various methods that can be used are designed to reduce the threat to domestic producers from other firms operating in the world economy.

🔗 **Time:** time is a key concept in relation to protectionism because some tools of protection, such as a quota or an embargo, could take effect immediately if required, whereas other tools, such as a subsidy, are likely to take a relatively longer period of time to have a significant impact on trade.

The different tools of protection and their impact

A number of different methods can be used to protect domestic industries:

Tool of protection	Impact
Tariffs	These are taxes that are imposed on imported products; the effect is to make the imported products more expensive than they would otherwise be and this should lead to a reduction in the demand for them, although the relative size of the actual effect will depend on the price elasticity of demand for the imported product.
Import duties	These operate in the same way as tariffs, raising the price of imported goods and so making them less likely to be demanded compared with alternative products produced in the domestic economy.
Import **quotas**	These are restrictions on the number or value of goods that can be imported, or the proportion of market share that they represent; the effect is to reduce the amount of imports purchased by domestic consumers and firms.
Subsidies	These are payments by a government to a domestic firm to help it keep down the costs of production; the effect is that, if the lower cost is passed on to consumers in the form of lower prices, domestic goods will be more competitively priced compared to imports, and demand for them is likely to increase.
Export subsidies	These are payments by a government to a domestic firm to help it keep down the costs of production of the products that it is intending to export; this is likely to increase the demand for the country's exports, if the demand for them is price elastic.
Exchange controls	These are restrictions on the buying and selling of foreign currency; the effect is to make it more difficult to finance the purchase of imported products.
Embargoes	These are complete bans on certain imported products, a decision that is usually taken for political, rather than economic, reasons; the effect is to make it impossible to purchase

	imported products from particular countries.
Excessive administrative burdens	This is where paperwork or 'red tape' is made more difficult; the effect is to make it much more difficult to get the imported products into the country.
Voluntary export restraints (VERs)	In some situations, one country might be fearful that another country to which it is exporting may decide to impose protectionist barriers on that trade, which could have dramatic consequences for the exporting country's economy. To reduce the likelihood of such import controls being established, the exporting country could establish a voluntary export restraint or restriction, limiting the amount that it will export. The importing country may then decide against imposing import controls on the exporting country.

- **tariff:** a tax that is imposed on an imported product to make it more expensive, in the hope that this will reduce demand for the product
- **import duty:** a duty that is imposed on an imported product to make it more expensive, in the hope that this will reduce demand for the product
- **quota:** a limit on the imported products that are allowed to enter a country; a quota can take the form of a limited quantity, a limited value or a limited market share
- **export subsidy:** a payment by a government to a domestic firm to help it keep down the costs of production of the products that it is intending to export, and thereby to increase demand for exports
- **exchange controls:** restrictions on the buying and selling of foreign currency, which make it more difficult to finance the purchase of imported products
- **embargo:** a ban on imports from particular countries, applied either to particular products or to all products from particular countries, usually for political, diplomatic or military reasons
- **voluntary export restraint (VER):** a decision, taken by an exporting country, to restrict its exports voluntarily in the hope that a country that it exports to will decide against imposing import controls

🔗 It is important that candidates can distinguish between a *tariff* and a *quota* as these two forms of trade protection are often confused in examination answers. A tariff refers to a tax or duty that is placed on an imported good. A quota is a restriction on the import of certain products, by quantity, value or market share.

The arguments for and against protectionism

The arguments for protectionism	The arguments against protectionism
• **Infant industry argument:** infant or **sunrise industries** need protection, at least as a temporary measure, to allow firms to be strong enough to compete with already established firms.	• **Trade diversion:** when trade barriers are established, it can lead to trade diversion and a certain amount of trade will be lost.

• **Declining** or **sunset industries:** these industries need to be protected, at least temporarily, to give time for the factors of production to be transferred to alternative uses. • **Strategic industries:** these industries, such as weapons production, may need to be protected because otherwise a country may be vulnerable to attack by an enemy. • **Anti-dumping measure:** firms that sell products that have been imported into a country may sell them not only cheaply to establish a market foothold, but at a price that is actually below the cost of production; this is known as **dumping** and a country may decide to use protectionist methods to protect itself from such dumping. • **Reducing a current account deficit:** a country may be experiencing a balance of payments deficit on current account (e.g. the value of its imports may exceed the value of its exports), and protectionism may be employed to try to overcome this deficit by restricting the imports coming into the country, although such a policy will not overcome the underlying reasons for the deficit. • **Raising revenue:** some protectionist methods, such as tariffs, raise revenue and so a government may use them for this reason.	• **Inefficiency:** protectionism can encourage industries to remain inefficient because they are protected from tough foreign competition, so there is an inefficient allocation of resources. • **Monopolies:** protectionism can allow monopolies to be created as foreign competition is reduced or eliminated. • **Distortion of markets:** protectionism involves a deadweight loss of consumer surplus and producer surplus, reducing economic welfare through higher prices and restricted consumer choice. • **Reduced world output:** trade barriers reduce world production. • **Quality of products:** protectionist barriers encourage consumers to buy domestically produced goods, but these may be of inferior quality.

- **infant industry argument:** the idea that a newly established industry should be given time to establish itself; it will, therefore, need to be protected, at least temporarily
- **sunrise industries:** industries that are new, or relatively new, and which are growing fast; it is expected that they will become very important in the future
- **sunset industries:** industries that have passed their peak and are now in decline, with no realistic hope of recovery
- **dumping:** the practice of selling a product at a price that is less than the cost of production

Protectionism can reduce the potential benefits of free trade and this is why it is strongly discouraged by the World Trade Organization.

🔗 Candidates often state that dumping is where a product is sold cheaply to gain access to a market. Where a product is sold at a price that is cheaper than that charged by domestic producers, this does not mean that it is an example of dumping. For dumping to take place, the product has to be sold at a price that fails to cover the marginal cost of production.

KEY SKILL

Evaluation: you should be able to evaluate the relative arguments for and against protectionism and come to a judgement as to whether it is an appropriate policy for a country to adopt.

6.3 Current account of the balance of payments

The components of the current account of the balance of payments

The **current account of the balance of payments** is made up of the following four parts:

» Trade in goods — this is the **balance of trade in goods account**, i.e. of **exports** and **imports** of goods (sometimes called 'the visible trade balance').
» Trade in services — this is the **balance of trade in services account**, i.e. of exported services and imported services (sometimes called 'the invisible trade balance').
» **Primary income** — this refers to net income flows, such as **net investment income** (e.g. dividends and interest).
» **Secondary income** — this refers to net **current transfers** (e.g. transfers of money by governments and individuals); in the case of governments, this could relate to international aid or contributions to international organisations, while in the case of individuals, it could relate to gifts or charitable donations.

- **current account of the balance of payments:** this comprises trade in goods, trade in services, primary income and secondary income
- **balance of trade in goods account:** the trade in goods (e.g. cars) between countries
- **exports:** goods and/or services that are produced domestically in one country and sold to other countries
- **imports:** goods and/or services that are produced in foreign countries and consumed by people in the domestic economy
- **balance of trade in services account:** the trade in services (e.g. banking) between countries
- **primary income:** the net flows of profits, interest and dividends from investments in other countries
- **net investment income**: the net income that relates to investments (e.g. dividends on shares or interest payments)
- **secondary income:** net payments where there is no exchange of a product, including government transfers of income and transfers of income by private individuals
- **current transfers:** the net payments by governments and private individuals (e.g. grants for overseas aid or charitable donations)

🔗 It is important that candidates demonstrate an awareness of the four elements of the current account of the balance of payments in their examination answers. Such knowledge would be particularly useful in relation to multiple choice questions.

The definition of equilibrium and disequilibrium in the current account of the balance of payments

deficit: a negative balance in the current account of the balance of payments when expenditure exceeds income
surplus: a positive balance in the current account of the balance of payments when income exceeds expenditure
external balance: the balance between receipts and payments in relation to international transactions between one country and other countries in the world

Equilibrium in the current account of the balance of payments refers to a situation where a country is experiencing neither a deficit nor a surplus over a period of time in its current account.

Disequilibrium in the current account of the balance of payments refers to a situation where a country is experiencing a deficit or surplus over a period of time in its current account.

» A **deficit** in the current account means that the money going out of a country through the current account is greater than the money coming into the country.
» **A surplus** in the current account means that the money coming into a country through the current account is greater than the money going out of the country.

The term 'deficit' or 'surplus' is therefore said to refer to the **external balance** of a country.

The calculation of the current account balance

The balance of trade in goods

The balance of trade was defined earlier as the difference in value of exports and imports of goods (sometimes called 'visibles') over a given period.

An example of a calculation is:

Value of exports	US$440 billion
Value of imports	−US$620 billion
Balance of trade in goods	−US$180 billion

The balance of trade in services

The balance of trade in services was defined earlier as the difference in value of exported services and imported services (sometimes called 'invisible inflows and invisible outflows') over a given period.

An example of a calculation is:

Value of exported services	US$550 billion
Value of imported services	−US$470 billion
Balance of trade in services	US$80 billion

The balance of trade in goods and services

The balance of trade in goods and services is the sum of the balance of trade in goods and the balance of trade in services.
An example of a calculation is:

Balance of trade in goods	−US$180 billion
Balance of trade in services	US$80 billion
Balance of trade in goods and services	−US$100 billion

The current account balance

The current account balance (CAB) was defined earlier as the sum of the balance of trade in goods, the balance of trade in services, the primary income and the secondary income.

An example of a calculation is:

Balance of trade in goods	−US$180 billion

Balance of trade in services	US$80 billion
Primary income	US$300 billion
Secondary income	−US$50 billion
Current account balance	US$150 billion

KEY SKILL

Numerical skills: you need to be able to calculate the current account balance from the balances that are included in it.

The causes of imbalances in the current account of the balance of payments

There are a number of reasons why a country might be experiencing a persistent imbalance or disequilibrium in the current account of the balance of payments.

Deficit

marginal propensity to import: the proportion of an increase in income that is spent on imported goods and services

In relation to a persistent deficit, possible reasons could include the following:

» The foreign exchange rate could be too high, causing exports to be more expensive than they would otherwise be. If demand for these exports is price elastic, this could have a significant effect on the current account of the balance of payments, contributing to a deficit in the current account.
» Consumers in a country could have begun to increase their demand for imported products —that is, as incomes have risen, there has been an increase in the **marginal propensity to import**. This could also lead to a deficit in the current account of the balance of payments
» There could be changes in consumer tastes and preferences within the country and/or abroad which reduce the demand for a country's exports and increase the demand for imports.
» Low competitive strength in world markets, perhaps as a result of relatively low levels of productivity or low levels of spending on research and development, will adversely affect a country's exports.

Surplus

In relation to a persistent surplus, possible reasons could include the following:

» Technological changes in methods of production in domestic industries may lead to lower costs, lower prices and an improvement in the quality of products.
» A tightening of import restrictions will make it more difficult to import products from other countries.
» A relatively low level of inflation in an economy, which makes exports more price competitive in world markets, may increase the demand for them.

🔗 **The margin and decision making:** the marginal propensity to import is another example of the importance of the margin in relation to decision making.

🔗 In examination questions on imbalances in the current account of the balance of payments, candidates tend to assume that the question is referring to a deficit, but disequilibrium can refer to either a deficit or a surplus. It is possible for a country to experience a persistent surplus, as has been the case with China.

The consequences of imbalances in the current account of the balance of payments

Consequences of a persistent current account deficit for the domestic economy

» There will be an increase in unemployment if there has been a decrease in the demand for exports and an increase in the demand for imports, creating a deficit in the current account.
» A persistent deficit in the current account could lead to a reduction in business confidence, resulting in a fall in the level of investment in the domestic economy.
» If corrective action is taken in an attempt to reduce, and hopefully eliminate, the deficit, consumers will either have a restricted range of imported products to choose from (if quotas have been introduced) or have to pay more for the imported products (if tariffs have been introduced).
» This increase in prices could have an inflationary effect on the economy.
» A current account deficit may lead to a depreciation in the value of a country's exchange rate, although this would help to restore a country's competitiveness in world markets.
» A deficit may be caused by short-term consumption rather than long-term investment, although a deficit can enable an economy to have a higher standard of living.

Consequences of a persistent current account surplus for the domestic economy

» It increases a country's net assets by the amount of the surplus.
» A relatively high level of exports could lead to an increase in employment in the export sector of the domestic economy.
» Lower import spending may mean consumers are spending more on domestic products rather than buying foreign ones, and this greater demand for domestic goods could increase domestic employment.

Consequences of a disequilibrium in the current account for the external economy

» If the disequilibrium is a deficit, there is likely to be a move towards greater protectionism in the international economy, reducing the extent of the benefits that would otherwise have been obtained from trade.
» If the disequilibrium is a surplus, a country will be able to accumulate foreign assets (e.g. China's investment in many African countries).
» It could cause less output and less employment in those countries experiencing a deficit in their current account balance.

KEY SKILL

Analysis: it is important to be able to analyse the consequences of imbalances in the current account of the balance of payments in terms of both the domestic and external economy and in terms of both a deficit and a surplus. For example, there are potential issues involved in a country experiencing a persistent surplus in the current account of the balance of payments, such as the fact that one country's surplus is another country's deficit, and the country experiencing the deficit could introduce protectionist measures that could have damaging effects on international trade.

6.4 Exchange rates

The definition of exchange rate

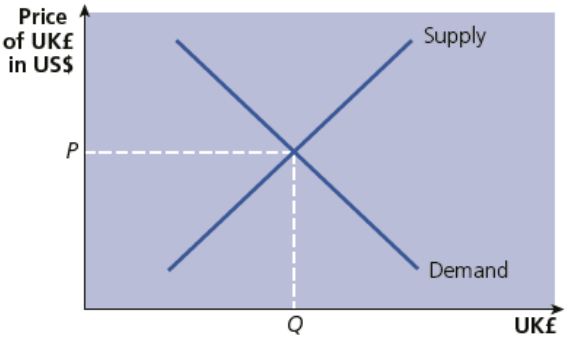
exchange rate: the value of one currency in terms of another

An **exchange rate** refers to the value of one currency in relation to another. It is the price of one particular currency expressed in terms of another.

The determination of a floating exchange rate

- **floating exchange rate:** an exchange rate that is determined, like any other free market price, by the market forces of demand and supply

There are a number of different systems in which exchange rates can be determined. One such mechanism is the floating exchange rate.



F6.2 price of sterling in US$

» A **floating exchange rate** is one that allows the value of a currency to be determined by the forces of demand and supply, just like the price of anything else in a free market.
» The vertical axis shows the price of UK£ in US$ and the horizontal axis shows the quantity of UK£.
» Equilibrium is established where the demand and supply curves intersect — that is, at a price of *P* and a quantity of *Q*.

Equilibrium and disequilibrium: in a floating exchange rate system, the price or exchange rate of a currency in terms of another currency will be determined by the equilibrium position, where the demand for the currency is equal to the supply.

KEY SKILL

Evaluation: you need to be able to compare and contrast the various advantages and disadvantages of a floating exchange rate and come to a judgement about such an exchange rate in particular circumstances.

The advantages and disadvantages of a floating exchange rate

Advantages	Disadvantages
• There is no need for a central bank to hold foreign currency reserves to use to intervene to maintain a particular exchange rate. • The value of an exchange rate will be an accurate price, determined by the demand for, and the supply of, the currency (i.e. the value adjusts automatically). • Changes in the exchange rate will reflect, and put right, disequilibrium in the balance of payments. For example, when there is a deficit, the exchange rate will depreciate, which should encourage exports and discourage imports. • A floating exchange rate means that interest rates can be set to meet domestic economic aims rather than to maintain a particular exchange rate value.	• Speculation, where buyers believe an exchange rate is wrongly valued and so buy or sell currency with the aim of making a profit, may affect the value of the exchange rate. • A floating exchange rate may be volatile, making economic planning and forecasting more difficult; this causes instability, which can discourage investment and trade. • A significant fall in the exchange rate can be a major cause of a rise in the rate of inflation.

The distinction between a depreciation and an appreciation of a floating exchange rate

- **depreciation:** a fall in the value of a floating exchange rate
- **appreciation:** a rise in the value of a floating exchange rate

It is important to distinguish between a depreciation and an appreciation of a floating exchange rate:

» **Depreciation** describes the situation when the value of a floating exchange rate goes down.
» **Appreciation** describes the situation when the value of a floating exchange rate goes up.

Causes of changes in a floating exchange rate

Changes in a country's floating exchange rate can be caused by a number of factors, including:

» the demand for the country's exports from other countries
» the demand for imports into the country
» inflation rates in different countries affecting international competitiveness
» perceptions of quality/reliability in relation to both exports and imports
» changes in interest rates in different countries, affecting movements of 'hot money'
» changes in average costs of production in various countries, affecting the relative level of competitiveness
» differences in the changes in relative technology between countries
» trends in tourism
» changes in levels of confidence in the economies of different countries and/or the possibility of speculation
» changes in a country's current account balance
» changes in the economic growth rates of different countries

AD/*AS* analysis of the impact of exchange rate changes

AD/*AS* analysis can be used to assess the impact of changes in the exchange rate on the domestic economy's equilibrium national income in relation to:

» the level of real output
» the price level
» the employment level

A country's exchange rate could be depreciated in order to encourage an increase in exports and/or a decrease in imports. Aggregate demand is made up of consumption (*C*), investment (*I*), government expenditure (*G*) and net exports (*X* − *M*). An increase in net exports, therefore, will lead to an increase in aggregate demand.

» In Chapter 4, section 4.3, it was pointed out that if there is an increase in aggregate demand, the *AD* curve will shift to the right.
» This can be seen in Figure 4.3 on page 67, where there is a shift in the aggregate demand curve from *AD*0 to *AD*1. This results in an increase in the level of real output and employment (from *Y*0 to *Y*1) and an increase in the price level (from *P*0 to *P*1).
» This effect, however, may not happen immediately. The reason for this is that buyers take time to adjust to price changes. A lower price for exports and a higher price for imports may have an effect on the level of aggregate demand over time, but the expected changes will not happen immediately.
» The depreciation may eventually have a positive effect on the external economy, if the purchase of exports is encouraged, but the eventual effect on the *domestic* economy may be negative because the cost of imported raw materials and component parts will have increased.

Another factor to take into account when assessing the effect of a change in the exchange rate on the domestic economy is in relation to the price elasticity of demand. A depreciation will make exports cheaper and imports dearer, but the ultimate effect of these price changes will depend on the price elasticity of demand for both the exports and the imports. If a depreciation is to be successful in terms of improving the current account of the balance of payments situation, the sum of the price elasticity of demand for exports and the price elasticity of demand for imports will need to be greater than 1.

🔗 **Equilibrium and disequilibrium:** *AD*/*AS* analysis can be used to assess the impact of exchange rate changes on a domestic economy's equilibrium national income.

6.5 Policies to correct imbalances in the current account of the balance of payments

Government policy objective of stability of the current account

» Three macroeconomic objectives of a government were identified in Chapter 5.
» It is now possible to add a fourth objective: the achievement of stability of the current account of the balance of payments over a period of time.
» The aim is for an equilibrium over time in this account so that deficits and surpluses in the current account of the balance of payments are approximately equal.
» However, such stability may not be possible, so a government may need to take action to correct an imbalance or disequilibrium in the current account, whether that is a deficit or a surplus.

The effect of fiscal, monetary, supply-side and protectionist policies on the current account

Correcting a persistent current account deficit

Disequilibrium in the current account of the balance of payments can come about through a persistent deficit. Government economic policies can be used to tackle this.

Fiscal policy

Deflationary fiscal policy (e.g. through an increase in taxation and/or a reduction in government expenditure) will create a downward multiplier effect in the economy, bringing down the level of aggregate demand. This is likely to reduce expenditure on imports.

Monetary policy

A deflationary monetary policy (e.g. through an increase in interest rates and/or a reduction in the money supply) is likely to reduce expenditure on imports. A government could also depreciate the exchange rate, making exports more competitive in world markets and imports less competitive in the domestic market.

Supply-side policy

Economists are agreed that the main way to reduce or eliminate a balance of payments deficit is to improve the quality of the goods produced in an economy, so that more people will buy them, both within the domestic market and abroad, and also to lower the price. Many economies have adopted a variety of supply-side policies to improve the competitiveness of their products. For example, privatisation and deregulation will increase competition in markets and make domestic firms more efficient, improving quality and lowering costs. Increased government spending on training and education could also lead to an increase in exports, reducing a current account deficit.

Protectionist policies

If a country is experiencing a deficit on the current account, protectionist policies can be used to reduce the value of imports entering a country. They can include:

» Tariffs
» Import quotas
» Export subsidies
» Embargoes
» Excessive administrative burdens (‘red tape’)

The main problem associated with such protectionist policies is that they involve restraints on free trade and are therefore generally opposed by the World Trade Organization.

Correcting a persistent current account surplus

Disequilibrium in the current account of the balance of payments can also come about through a persistent surplus. The following policy measures could be adopted.

Fiscal policy

If a country is experiencing a surplus in the current account, the government could reduce income tax, which would increase disposable income. This is likely to increase expenditure on imports and this would reduce the size of the surplus in the current account.

Monetary policy

The government could tackle the surplus by increasing the growth of the money supply and/or decreasing interest rates. This is likely to increase expenditure on imports. It could also appreciate the exchange rate, making exports less competitive in world markets and imports more competitive in the domestic market. These measures are likely to reduce the size of the surplus in the current account, possibly eliminating it altogether.

Reduced protectionism

If protectionist methods are reduced or eliminated entirely, this is likely to increase expenditure on imports and this will reduce the size of the surplus in the current account.

A level:

7 The price system and the microeconomy

7.1 Utility

The definition and calculation of total utility and marginal utility

- **total utility:** the total amount of satisfaction obtained from the purchase of a number of units of a product
- **marginal utility:** the increase in utility that a consumer gains from the consumption of an extra unit of a product

The term 'utility' refers to the satisfaction that is derived from the consumption of a particular product. It is possible to distinguish between total utility and marginal utility:

» **Total utility** is the total satisfaction obtained from the consumption of a given number of units of a particular product. It is calculated by adding together all the utility or satisfaction derived from the consumption of a given number of goods or services. The formula is:

$TU = U1 + MU2 + MU3$ etc.

It is therefore the utility gained from the first unit plus the marginal utility gained from the second unit plus the marginal utility gained from the third unit and so on.

» **Marginal utility** is the increase in utility that a consumer gains from consuming an additional unit of a product. It is calculated by measuring the increase that a consumer gains from consuming an extra unit of a good or service. The formula is the difference in the total utility divided by the number of goods or services.

Diminishing marginal utility

- **diminishing marginal utility:** the principle that the marginal utility of consuming successive units of the same product will fall

» The law or principle of **diminishing marginal utility** states that the consumption of successive units of a product will eventually lead to a fall in marginal utility — that is, as a person consumes more units of a product, the satisfaction provided by each unit will be progressively less and less.
» The impact of diminishing marginal utility on total utility is that total utility will also start declining.
» Maximum total utility is achieved when marginal utility is zero because if marginal utility is zero, this adds nothing to total utility.

The equi-marginal principle

equi-marginal principle: a consumer will maximise total satisfaction by equating the utility or satisfaction per unit of money spent on the marginal unit of each product consumed

The **equi-marginal principle** is an example of consumer microeconomic equilibrium and shows the relationship between the marginal utility obtained from the consumption of different products and the prices paid for those products. It can be represented in the following way:

$$\frac{MU_a}{P_a} = \frac{MU_b}{P_b} = \frac{MU_c}{P_c}$$

To maximise their utility or satisfaction, it is assumed that consumers will consume up to the point shown above — that is, the extra satisfaction, in relation to the money spent, on the last unit of product A will equal the extra satisfaction, in relation to the money spent, on the last unit of product B and so on.

🔗 **The margin and decision making:** the importance of the margin as a key concept in economics can be clearly seen in relation to diminishing marginal utility and the equi-marginal principle. This is because a consumer will be in equilibrium, assuming a given level of income, when it is not possible to switch expenditure from one product to another to increase total utility.

The derivation of an individual demand schedule

There is a relationship between the law of diminishing marginal utility and the derivation of an individual demand schedule and curve. If the marginal utility of consuming an extra item of a product continually falls, a consumer will be unwilling to pay as much for each successive unit consumed. This explains why a demand curve is downward sloping from left to right, showing that there is a relationship between changes in price and changes in the quantity demanded.

The limitations of marginal utility theory and its assumptions of rational behaviour

The law of diminishing marginal utility is based on a number of assumptions, and if these assumptions do not apply, there are clear limitations to the theory.

The assumptions include the following:

» The idea of utility or satisfaction that a consumer gains from the consumption of particular items of a product — but this assumes that satisfaction can be easily measured. Utils are a rather abstract unit of measurement.
» The idea that consumers behave in a rational way — but is this always the case? Advertising can impact on consumer behaviour and distort consumer choices, so that consumers do not act in a rational way.
» The idea that consumers have limited incomes — but it is possible that incomes will rise significantly over a period of time (although incomes would still be limited at any particular time).
» The idea that consumers aim to maximise their total utility — but is this always going to be the case? For example, behavioural economics suggests that consumers may act on the basis of impulse or emotion and not on the basis of utility maximisation.
» The idea that prices are constant — but it may be the case that the prices of products are continually changing.
» The idea that consumer tastes and preferences remain constant — but this may not always be the case.

KEY SKILLS

Evaluation: you need to be able to criticise the validity and usefulness of the theory of marginal utility.

Analysis: you need to be able to critically analyse marginal utility theory and its assumptions of rational behaviour, demonstrating an understanding of its limitations — for example, in analysing the differences between rational behaviour and the ideas of behavioural economists.

Paradox of value

- **paradox of value:** the fact that certain products that are essential to survival (e.g. water) are cheaper than products that are less important to survival (e.g. diamonds)

- **behavioural economics:** an approach to decision making which argues that the behaviour of individuals is often based on ideas that do not correspond to the traditional view of rational economic behaviour

» Certain products that are vital to our survival, such as water, are not as expensive as products that are less crucial, such as diamonds. This is known as the **paradox of value**.
» The explanation of this paradox can be seen in terms of marginal utility and total utility.
» For example, people will consume water up to the point where marginal utility is zero and total utility is maximised, whereas they will demand diamonds where marginal utility is high, but total utility is low.

🔗 Candidates need to show they understand why diamonds are more expensive than water, illustrating the concept of paradox of value. Water is generally abundant and so its marginal utility, and therefore its price, is relatively low.

Rational behaviour versus behavioural economic model

A major assumption that underpins marginal utility theory is that consumers can always be expected to act rationally.

Rational behaviour essentially means the following:

» Individuals will take decisions to maximise their own utility or satisfaction.
» Individuals have access to all the information that they need to make a decision at zero cost.
» Individuals take decisions that are based on a very careful comparison of the benefits and costs to achieve the optimum outcome.
» Decisions will be taken by individuals based on changes at the margin.
» The preferences of individuals and their attitude to risk are assumed to be fixed.

Behavioural economics, on the other hand, stresses that the behaviour of individuals may differ greatly from these assumptions — in other words, it often appears to be ‘irrational’ rather than rational.

Behavioural economics attempts to explain such behaviour. It is not limited to economics, but brings in concepts and theories from sociology and psychology.

It is assumed that individuals take optimal decisions based on the information that is available to them, but behavioural economics stresses that rational behaviour should not always be expected:

» Behavioural economists suggest that there may be so much information available that individuals often take the same decision that they always have — that is, they often base decisions on past experience.
» In this sense, individuals are ‘creatures of habit’ and are strongly influenced by brands and the uses made of branding in marketing.
» The individual’s ability, and indeed willingness, to act in a rational manner is therefore restricted in certain ways.

🔗 Candidates need to show they understand that behavioural economics suggests that individuals may take decisions for reasons other than those that traditional economics has stressed are rational.

KEY SKILL

Evaluation: you need to be able to consider to what extent consumers act in a rational way when taking decisions about expenditure, and are thus able to maximise their consumer welfare and reach consumer equilibrium as predicted by marginal utility theory.

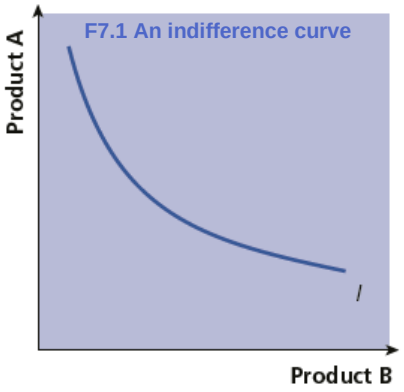
7.2 Indifference curves and budget lines

The meaning of an indifference curve and a budget line

Indifference curve

- **indifference curve:** a curve showing all the possible combinations of two products between which an individual consumer is indifferent

Consumer preferences can be represented by an **indifference curve**. An indifference curve shows all the possible combinations of two products between which a consumer is indifferent —that is, the various combinations that provide equal satisfaction or utility.



» The indifference curve slopes downwards from left to right.
» It is convex to the origin.
» No two indifference curves can ever intersect.
» The slope of an indifference curve shows the marginal rate of substitution — that is, the number of one good that an individual is prepared to give up in order to obtain additional items of the other good.

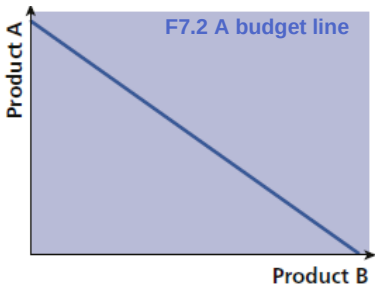
🔗 **The margin and decision making:** the slope of an indifference curve shows the marginal rate of substitution — that is, the number of one good an individual is prepared to give up in order to obtain additional items of the other good.

KEY SKILL

Diagrams: in drawing indifference curves, it must be remembered that no two indifference curves can ever intersect.

Budget line

budget line: a line showing all the possible combinations of two products that a consumer would be able to purchase with fixed prices and a given income; sometimes known as a 'consumption possibility line'



A budget line shows how consumers are constrained in terms of what they are able to buy as a result of their limited income and the prices of the goods and services.

A budget line has the following characteristics:

» It shows the possible combinations of two products that a consumer is able to purchase with a given income and fixed prices.
» Each of the combinations would cost the same total amount.
» Any point along the line will show the maximisation of consumption at the given income level.

Figure 7.2 shows all the combination of products A and B that can be purchased by a consumer, assuming that the income of the consumer and the prices of the products remain fixed.

The causes of a shift in the budget line

Figure 7.2 assumes that the income of the consumer and the prices of the products remain fixed. However, shifts in the budget line can occur because of changes in income and changes in price:

» **Change in income**: if there is an increase in income (assuming ceteris paribus), a consumer will be able to spend more on both products. If the prices of the two products remain unchanged, the budget line will shift outwards parallel to the original budget line.
» **Change in price**: if there is a change in the price (assuming ceteris paribus) of one of the products, the budget line will not shift parallel to the right, but will pivot. For example, if the price of product A remains the same, but the price of product B falls, then a consumer will buy the same amount of product A, but more of product B.

🔗 You need to understand the difference between a parallel shift and a pivot of a budget line. If the prices of the two products remain the same, the budget line will shift parallel to the original budget line. However, if there is a change in the price of one of the products, the budget line will pivot. Also, you need to consider relative prices; if the price of A *and* the price of B fall, but by different percentages, the slope of the budget line will also change.

The income, substitution and price effects for normal, inferior and Giffen goods

Price effect

- **price effect:** the effect on the consumption of a product that occurs as a result of a price change

A **price effect** refers to a change in the consumption of a product as a result of a change in its price. A change in the price of a product can actually bring about two effects:

» an income effect
» a substitution effect

Income effect

- **income effect:** the effect on consumption of a change in real income that occurs as a result of a price change

An **income effect** refers to the situation where a change in the price of a product will bring about a change in real income. Although there is no change in nominal income, a rise or fall in the price of a product will have a real income effect — a person will be able to buy more or less of a product, and other products, as a result of the change in price of the product.

Substitution effect

substitution effect: the effect of a rise or fall in the price of a product on the utility or satisfaction obtained from each unit of money spent on that product; as a result of the price change, expenditure can be rearranged to maximise the utility or satisfaction gained

The second effect of a change in the price of a product relates to the utility or satisfaction obtained from the consumption of a product. A **substitution effect** means that a rational consumer will substitute in favour of a product that has now become relatively cheaper.

Normal, inferior and Giffen goods

- **Giffen good:** a good where a higher price causes an increase in demand, reversing the usual law of demand

It is important to understand how the price effect can apply to normal, inferior and Giffen goods:

» **Normal good:** with a normal good, if its price falls, the substitution and income effects work in the same direction — there will be increased consumption of the good.
» **Inferior good:** with an inferior good, if its price falls, the substitution and income effects work in opposite directions. However, there will be increased consumption of the good because the substitution effect is greater than the income effect, leading to an overall positive price effect.
» **Giffen good:** with a Giffen good, a product for which demand increases as the price increases and demand decreases as the price decreases, if its price falls, the substitution and income effects work in opposite directions. However, there will be reduced consumption of the good because the income effect is greater than the substitution effect.

🔗 It is important to remember that the income effect can be positive or negative and this is what determines whether a good is a normal good or an inferior good.

KEY SKILL

Application: A Giffen good is a good for which demand rises as its price rises, thus contradicting the law of demand. It is named after the Scottish economist, Sir Robert Giffen (1837–1910), who observed poor people buying more bread as its price rose. Whereas the demand for a normal good falls as its price rises, the demand for a Giffen good rises as its price rises, such as with basic foods like rice and wheat.

Limitations of the model of indifference curves

There are a number of limitations of the model of indifference curves, including the following:

» It oversimplifies the situation in relation to consumer behaviour and consumer choices, as indifference curve analysis is based on a two-product model.
» It makes unrealistic assumptions about human behaviour, assuming that consumer behaviour is always rational.
» It is incompatible with the reality of economic action, which demonstrates preference rather than indifference.
» Consumer preferences may change in time, making specific indifference curves less relevant.

KEY SKILL

Analysis: you need to be able to judge the strengths and weaknesses of the various limitations of the model of indifference curves.

7.3 Efficiency and market failure

- **efficient resource allocation:** the optimal use of scarce inputs to produce the largest possible output

The term ‘efficiency’ generally means **efficient resource allocation** — using resources in the most economical way possible, or making maximum use of the resources that are available. It can also refer to the idea of obtaining the maximum output for given inputs.

The concept of economic efficiency can be divided into two types:

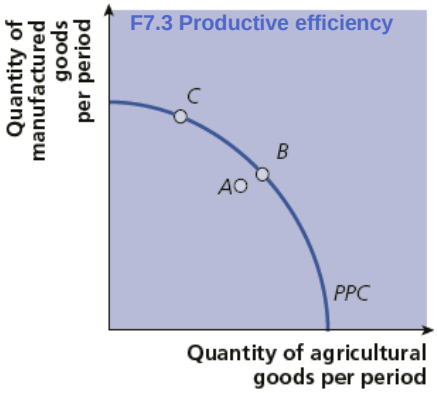
» Productive efficiency
» Allocative efficiency

Definitions of productive efficiency and allocative efficiency

- **productive efficiency:** where a firm operates at the minimum of its average cost curve; or where a whole economy is operating on its production possibility curve
- **average cost:** the total cost of employing all the factor inputs divided by the number of units produced; also known as 'average total cost'
- **technical efficiency:** where a firm produces the maximum output possible from given inputs, shown by the lowest point on the lowest possible average cost curve
- **cost efficiency**: where a firm uses the most appropriate combination of inputs of factors of production, given the relative costs of those factors
- **allocative efficiency:** a situation that describes the extent to which the allocation of resources in an economy matches consumer preferences and *P = MC*

Productive efficiency

One way of measuring **productive efficiency** is in terms of the **average cost** (or average total cost) of production. Productive efficiency can be defined as the minimum average cost at which output can be produced.



Productive efficiency involves two elements:

» **Technical efficiency** is where the best possible use is made of the inputs, or factors of production, used in the production process. This means that as much output as possible is produced from given inputs.
» **Cost efficiency** relates to whether the best set of inputs has been used in the production process; this will be influenced by the relative costs of different inputs, such as labour and capital.

Allocative efficiency

The most efficient allocation of resources in an economy will be the one that fits the needs and wants of consumers most closely. The more that firms in an economy can respond effectively to changes in the demand of consumers, the closer the economy can get to a situation of **allocative efficiency**.

The conditions for productive efficiency and allocative efficiency

Productive efficiency

As well as being seen in the micro context of a particular firm, the concept of productive efficiency can be seen in the macro context of the whole economy, as shown in Figure 7.3.

» An economy can decide between the production of manufactured goods (shown on the vertical axis) and the production of agricultural goods (shown on the horizontal axis). It shows the combination of these two types of goods in an economy.
» Productive efficiency is not taking place at point *A*, inside the *PPC*, because by moving to a point on the *PPC* it would be possible
for the economy to produce more of both goods.
» Any point on the *PPC* will indicate productive efficiency — points *B* and *C* are both productively efficient.
» There is a trade-off between points on the *PPC* in relation to the production of manufactured goods and the production of agricultural goods. For example, at point *C*, more manufactured goods will be produced in the economy than at point *B*, but fewer agricultural goods will be produced.

» Both points *B* and *C* are productively efficient, but it is not possible to make a judgement as to whether the society is better off at point *B* or *C* without knowing the preferences of the consumers in the society. It is therefore necessary to consider another type of efficiency, allocative efficiency.

🔗 Productive efficiency has been defined as where a firm operates at the minimum of its average cost curve, but this does not necessarily have to be at the minimum point of the long-run average total cost curve. It can be at any point of production along that curve.

Allocative efficiency

» Allocative efficiency can be clearly seen in the relationship between price and marginal cost.
» The marginal cost of production is the cost of producing one more product.
» When this cost is equal to the price charged for the product, there is a situation of allocative efficiency. This is because the value that is put on the resources used to produce the product by the producer (*MC*) is equal to the value put on the product by the consumer (price). If a greater or lesser amount of the product were produced, price and marginal cost would no longer be equal.

🔗 In a question on economic efficiency, it is important that both productive efficiency and allocative efficiency are considered. This is because economic efficiency is defined as the maximum number of goods and services that can be produced with a given amount of inputs, and for this situation to exist, the scarce resources in an economy need to be both productively and allocatively efficient.

🔗 **The margin and decision making:** another example of the importance of the concept of the margin in economics is in relation to allocative efficiency, where decisions are taken in terms of how scarce resources in an economy are used, and this is where price is equal to marginal cost.

Pareto optimality

Optimum resource allocation

- **optimum resource allocation:** the best allocation of resources possible in the situation of scarcity

» The word ‘optimum’ means the best possible outcome in a given economic situation.
» **Optimum resource allocation** refers to a situation where the best possible allocation of scarce resources exists. This is when both productive efficiency and allocative efficiency occur.
» The idea of an optimum, or optimality, is an important concept in economics, given that resources are scarce; when resources are allocated in an optimum way, it means that they are used in the most efficient way possible.

🔗 **Scarcity and choice:** the idea of optimal resource allocation is very important given that resources are scarce. When choices are made so that these scarce resources are allocated in an optimum way, the effect is that they are allocated in the most efficient way possible.

🔗 In examination questions on the concept of efficiency, candidates should ensure the idea of optimality is included in their answer, emphasising that this means the best possible allocation of scarce resources in a given situation.

The meaning of Pareto optimality

- **Pareto optimality:** a particular use of the term ‘optimality’ associated with the Italian economist Vilfredo Pareto, who stated that this situation existed when it was not possible to reallocate resources to make someone better off without making someone else worse off

» Vilfredo Pareto (1848–1923), an Italian economist, argued that the best possible allocation of scarce resources existed when it was impossible to make one person better off without making another person worse off. This is known as **Pareto optimality**.
» In such a situation, there must be an optimal allocation of resources — that is, the inputs or resources are used in the most efficient possible way (productive efficiency) and the output produced by the resources provides the maximum possible utility or satisfaction to consumers (allocative efficiency).
» An improvement in optimality can only occur when one person is made better off without making anyone else worse off.
» It can be shown using a production possibility curve as all points on a *PPC* are Pareto optimal.

🔗 It is important that candidates can demonstrate they understand what is meant by Pareto optimality — the idea that it is impossible to make one person better off without making another person worse off.

The definition of dynamic efficiency

- **dynamic efficiency:** the greater efficiency that can result from improvements in technical or productive efficiency over a period of time

Productive efficiency and allocative efficiency can both be considered as examples of *static* efficiency. This means that they are concerned with the allocation of scarce resources at a particular moment in time.

Dynamic efficiency, on the other hand, is concerned with changes in the allocation of scarce resources over a period of time. It can be shown by the downward movement of the *LRAC* curve. It can occur for a number of reasons:

» **Product innovation:** as a result of new products appearing through research and development, invention and innovation.
» **Process innovation:** it is also possible for new methods of production to be introduced as a result of developments in technology.
» **New techniques of management:** there could also be changes in techniques of management as a result of investment in human capital.

🔗 It is important that candidates can demonstrate they understand the difference between static efficiency, as in the case of productive and allocative efficiency, and dynamic efficiency, which takes place over a period of time.

The definition of market failure

- **market failure:** a market imperfection which gives rise to an allocation of scarce resources that is not as efficient as it might have otherwise been

Market failure can be defined as a market imperfection which gives rise to an allocation of scarce resources that is not as efficient as it might have otherwise been. It therefore occurs when there is an inefficient allocation of resources in a free market.

The main reasons for market failure:

Type of market failure	Explanation
Merit goods	These are goods that are regarded as socially desirable and that would be underproduced and underconsumed in a free market.
Demerit goods	These are goods that are regarded as socially undesirable and that would be overproduced and overconsumed in a free market.
Public goods	These are non-rival and non-excludable goods that would not be provided at all in a free market because it would be impossible to charge a price for them.
Externalities	Externalities are costs and benefits that affect third parties. Negative externalities may result from the overproduction and/or overconsumption of a good or service and positive externalities may result from the underproduction and/or underconsumption of a good or service.
Information failure	There may be a lack of full information to make an informed decision, so the allocation of resources may not be as efficient as it otherwise would be. The existence of asymmetric information is a particular form of market failure; this exists when one party has much more information than another party and uses that information advantage to exploit the other party.
Imperfect competition	This occurs when there is a departure from the situation of perfect competition. There can be different market imperfections and one of these is the existence of monopolistic elements in an economy, where one firm can have monopoly power in a market.
Inequality in the distribution of income and wealth	A free market may lead to a very unequal distribution of income and wealth, giving some people more influence in a market than others.
Factor immobility	In a perfect market, the factors of production would be able to move easily from one market to another, but the existence of occupational and geographical immobility of factors means that this does not always happen.
Price instability	This occurs where prices can change rapidly in a short period of time. In some markets, there can be a great deal of price instability, especially in agricultural markets.

🔗 **Efficiency and inefficiency:** you need to understand the importance of efficiency in relation to the different types (i.e. productive efficiency, allocative efficiency and dynamic efficiency), and you also need to be aware of the existence of inefficiency in relation to different examples of market failure.

KEY SKILL

Application: you need to be able to apply the idea of market failure to a variety of different situations, such as in relation to the non-existence of public goods, the under provision of merit goods and the overprovision of demerit goods.

7.4 Private costs and benefits, externalities and social costs and benefits

The definition and calculation of social costs as the sum of private costs and external costs, including marginal social costs, marginal private costs and marginal external costs

- **social costs:** the sum of private costs and external costs (i.e. the total cost to society of an economic decision or activity)
- **external costs:** the negative effects imposed on third parties not involved in an economic decision or activity and not compensated for

Social costs (*SC*) represent the true cost to society — in other words, they include not only private costs (*PC*), but also the **external costs** (*EC*) imposed on the whole society as a result of an economic activity.

🔗 Candidates need to demonstrate that they understand that social costs include *both* private costs *and* external costs. A common mistake is to confuse social costs and external costs.

The definition and calculation of social benefits as the sum of private benefits and external benefits, including marginal social benefits, marginal private benefits and marginal external benefits

- **social benefits:** the sum of private benefits and external benefits (i.e. the total benefit to society of an economic decision or activity)
- **external benefits:** the positive effects gained by third parties not involved in the economic decision or activity and not paid for

Social benefits (*SB*) represent the true benefit of an economic activity to society — that is, they include not only private benefits (*PB*), but also the **external benefits** (*EB*) that are advantageous to the whole society as a result of an action. An example of an external benefit is the additional income that is created in an area when a new business locates there. This will lead to an increase in spending in the community.

🔗 Candidates need to demonstrate that they understand that social benefits include *both* private benefits *and* external benefits. A common mistake is to confuse social benefits and external benefits.

The definition of positive externality and negative externality

Externality

- **externality:** a cost or benefit of either consumption or production that is paid for or enjoyed not by the consumer or the producer, but by a third party

- **third party:** individuals or groups that are not the main parties in a transaction, but are still affected by it
- **spill over effect**: the effect of certain decisions that have an impact on third parties (i.e. those who are neither the producers nor the consumers of a particular product)

An **externality** has certain characteristics:

» It arises if a **third party** is affected by the actions and behaviour of others.
» A third party can be regarded as someone who is not directly involved in an economic activity.
» An externality can also be described as a situation where there is a **spillover effect**.
» An externality is external to a market transaction and so is not reflected in market prices.
» It can occur in relation to consumption or production.
» It is a form of market failure because if the cost or benefit is not reflected in market prices, it cannot be taken into account by all the parties involved in a transaction; that is why the costs or benefits that result from such a transaction affect a third party that is not directly involved in the transaction.

🔗 Candidates need to understand what is meant by a 'third party' — that is, someone who is not directly involved in the consumption and/or the production of a product. The third party, however, can still be affected in different ways by consumption and/ or production decisions taken by others. For example, in the case of passive smoking, a person who is not actually smoking a cigarette can be badly affected by the actions of others around them. Similarly, in the case of drink-driving, a motorist may be sober and driving very well, but may be involved in an accident as a result of the poor driving of another motorist under the influence of alcohol.

Positive externality

- **positive externality:** the external benefit that may occur as a result of an action, bringing some benefit to a third party

A **positive externality** refers to the benefit that can be gained by a third party — that is, someone who is not directly involved in a transaction. A positive externality can be seen in relation to either consumption or production.

Negative externality

- **negative externality:** the external cost that may occur as a result of an action, bringing some disadvantage to a third party

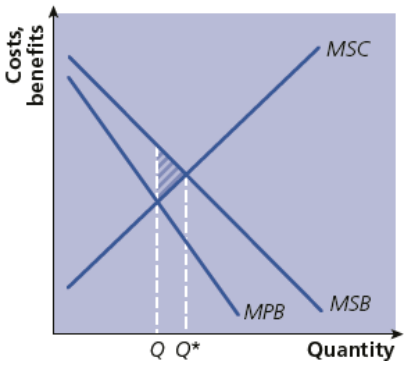
A **negative externality** refers to the disadvantage that can affect a third party — that is, someone who is not directly involved in a transaction. Like a positive externality, a negative externality can be seen in relation to either consumption or production. For example, a third party might be negatively affected by the consumption of a loud music system by a neighbour. In terms of production, a third party might be negatively affected by pollution caused by a factory in the neighbourhood.

Positive and negative externalities of both consumption and production

Positive consumption externalities

- **positive consumption externality:** a third-party effect that influences the consumption side of a market in a positive or beneficial way
- **welfare gain:** a situation that arises when the marginal social benefit exceeds the marginal social cost, leading to a socially efficient allocation of resources

» A **positive consumption** externality is where there is a beneficial spill over effect on a third party arising from the consumption of a good or service — for example, the enjoyment provided by the views of private gardens.
» This can be seen in Figure 7.4. *MPB* shows the marginal private benefit, but *MSB* shows the marginal social benefit.
» The equilibrium which takes into account the marginal social benefit and the marginal social cost (i.e. the socially optimum equilibrium) would be at quantity *Q**, a greater output than at *Q* where only the marginal private benefit is being taken into account.
» The **welfare gain**, resulting from increasing output to the socially optimum equilibrium, where *MSC* crosses *MSB*, rather than *MPB*, is shown by the shaded triangle.

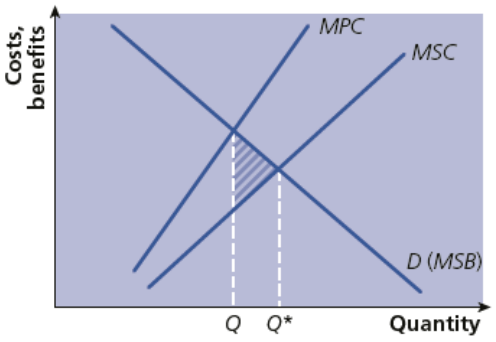


F7.4 A positive consumption externality

Positive production externalities

- **positive production externality**: a third party effect that influences the production side of a market in a positive or beneficial way

» A **positive production externality** is where there is a beneficial spill over effect on a third party arising from the production of a good or service.
» For example, if a firm purifies its waste water, this might be beneficial to a local fish farm in the area.
» This can be seen in Figure 7.5. *MPC* shows a firm's marginal private cost, but *MSC* shows the marginal social cost, which is lower than the *MPC*. The equilibrium taking account of the positive production externality is at output *Q** where *MSC* = *MSB* (the socially optimum equilibrium). This output at *Q** is greater than at *Q*. The welfare gain resulting from the extra output is shown by the shaded triangle.

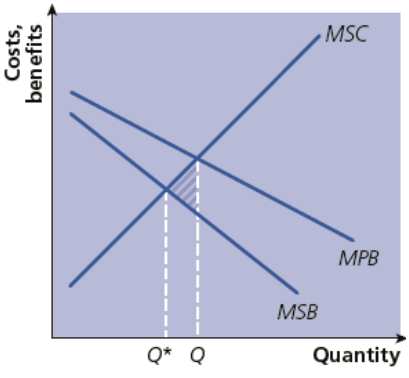


F7.5 A positive production externality

Negative consumption externalities

- **negative consumption externality:** a third-party effect that influences the consumption side of a market in a negative or disadvantageous way
- **welfare loss:** a situation that arises when the marginal social cost exceeds the marginal social benefit, leading to a socially inefficient allocation of resources

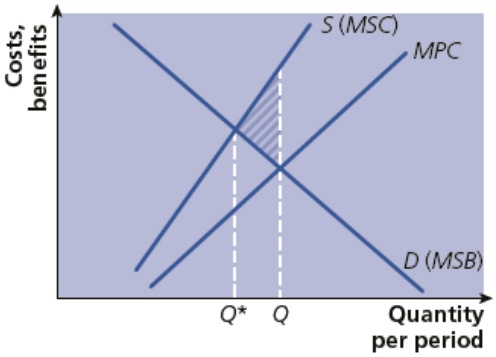
» A **negative consumption externality** is where there is a negative spill over effect on a third party arising from the consumption — for example, a person playing loud music that irritates the neighbours living nearby.
» This can be seen in Figure 7.6. An individual's appreciation of the music is shown by the marginal private benefit, *MPB*, but the negative effect on the neighbours means that the social marginal benefit, *MSB*, is lower.
» The equilibrium position for the person playing the music is at *Q*, where *MPB* and *MSC* cross, but the equilibrium position for the whole community, taking into account the negative effect of the loud music on the neighbours, is at *Q** where *MSC* and *MSB* intersect (the socially optimum equilibrium). The **welfare loss** of the additional output is shown by the shaded triangle.



F7.6 A negative consumption externality

Negative production externalities

- **negative production externality:** an externality that affects the production side of a market in a negative or disadvantageous way



F7.7 A negative production externality

» A **negative production externality** is where there is a negative spillover effect on a third party arising from the production of a good or service — for example, a factory producing noise and/or air pollution.
» This can be seen in Figure 7.7. on the right. The marginal private cost, *MPC*, only takes into account the cost of production to the firm. The marginal social cost, *MSC*, takes into account the full social costs of the production.
» For the firm, the equilibrium position would be at *Q*, but for society as a whole the equilibrium position would be at *Q** where *MSC* and *MSB* intersect (the socially optimum equilibrium).
If output was at *Q* rather than *Q**, there would be a welfare loss shown by the shaded triangle.

🔗 **the margin and decision making:** the concept of the margin is very important in relation to decisions that bring about externalities in terms of marginal private benefit, marginal social benefit, marginal private cost and marginal social cost.

Deadweight welfare losses arising from positive and negative externalities

- **deadweight loss:** the loss of economic efficiency that occurs when the socially optimal quantity of a product is not produced

» **Deadweight loss** is a measure of lost economic efficiency when the socially optimal quantity of a product is not produced.
» Non-optimal production can be caused by a positive or a negative externality.
» Externalities bring about a deadweight loss as a result of the differences between marginal social cost or benefit and marginal private cost or benefit.
» The deadweight loss, or welfare loss as it can also be called, is the decreased economic well-being caused by the existence of negative externalities. It is shown by the shaded triangle in Figure 7.6 (in the case of a negative consumption externality) and in Figure 7.7 (in the case of a negative production externality).
» There can also be a welfare gain in relation to positive externalities. It is shown by the shaded triangle in Figure 7.4 (in the case of a positive consumption externality) and in Figure 7.5 (in the case of a positive production externality).

Asymmetric information and moral hazard

Asymmetric information

- **Asymmetric information:** a situation in which there is unequal knowledge between the parties of a transaction, resulting in an advantage to the party with additional knowledge

Asymmetric information is a form of market failure that exists when one individual or party has much more information than another individual or party and uses that information advantage to exploit the other party. Information asymmetry therefore creates an imbalance of power — for example, when the seller of a product, such as a used/second-hand car, knows more about the good or service than the buyer.

Moral hazard

- **moral hazard:** a situation in which a person takes a decision about how much risk to take in the knowledge that someone else bears the cost of that risk

» **Moral hazard** occurs when someone increases their exposure to risk because someone else bears the cost of those risks.
» It therefore refers to a situation where an individual has an incentive to alter their behaviour when the potential risk is borne by others.
» For example, if a government promises to support businesses that are operating at a loss, such as at a time of economic recession, it can encourage those businesses to take greater risks.
» Another example is in relation to insurance when a person is more willing to take a risk in the knowledge that this will be covered by the insurance policy.

The use of costs and benefits in analysing decisions

- **cost–benefit analysis:** an analysis of a project which includes a valuation of the total costs and total benefits involved, including private and external costs and private and external benefits

Cost–benefit analysis provides a framework where all the costs and benefits of an investment project can be analysed, not only private ones. It is a particular feature of major transportation projects.

the four key stages of a cost–benefit analysis:

Stage	Description
1 The identification of all the relevant costs and benefits involved in the project	The first stage involves identifying all relevant costs and benefits: the private costs, the private benefits, the external costs and the external benefits.
2 Deciding the monetary value of all the relevant costs and benefits involved in the project	The second stage involves putting a monetary value on all of these costs and benefits, including those where a market price can be established and those where a market price is not easily established, such as placing a value on time.
3 The forecasting of the future costs and benefits involved in the project	The third stage involves forecasting the costs and benefits of an investment project into the future; this forecasting will be based on estimates of future costs and benefits.
4 The interpretation of the results of the cost–benefit analysis so that an appropriate decision can be taken	The fourth stage involves the compilation of all the data obtained as a result of the cost–benefit analysis and using them to ensure that the decision-making process is an informed one.

The key features of cost–benefit analysis are as follows:

» It takes the full social costs and the full social benefits into account before a decision is made as to whether a particular investment project should go ahead.
» It takes a wide view of a project, considering its full impact on an economy and society over a period of time.

🔗 In an answer to a question on cost–benefit analysis, candidates should emphasise that the advantage of this approach is that it can take into account the full social costs and social benefits involved in an investment project.

KEY SKILL

Application: cost–benefit analysis can be used whenever it is necessary to compare the advantages and disadvantages of a project, such as the building of a road, an airport runway or a railway line.

The problems facing the process of cost–benefit analysis:

Problem	Explanation
Absence of market prices	Cost–benefit analysis takes into account all costs and benefits, but not all of these will have a market price. For example, it is hard, if not impossible, to establish the price of pollution.
Need to use shadow prices	In the absence of market prices, values will need to be estimated through the use of shadow prices (i.e. prices that are estimated rather than accurately calculated), but this will not always be accurate (e.g. the valuation of time or an accident).
The future	The costs and benefits of any project, such as the building of a road or a runway, will be spread over a long period of time, so it will be difficult to compare the interests of present and future generations.
Spillover effects	It is recognised that a project will have spillover effects, but it may be difficult to establish what these are and how far they extend. For example, the external costs of pollution of building an airport are likely to be concentrated around the airport, while the benefits of building the airport can be more widely dispersed, such as the benefits of increased tourist flow.
Political versus economic decisions	Cost–benefit analysis may produce a decision in favour of a particular project, but it may still be rejected because of political decisions.

🔗 Knowledge of net present value is not required in answers to questions on cost–benefit analysis.

KEY SKILL

Application: there are many examples of investment projects that have used cost–benefit analysis, including the Hong Kong to Macau Bridge in China, Bengaluru Airport in India, the

London Underground Victoria Line in England, California High-Speed Rail in the USA and the provision of solar power in China.

7.5 Types of cost, revenue and profit, short-run and long-run production

The short-run production function

Fixed and variable factors of production

- **fixed factors of production:** resource inputs that exist in the short run when the quantity of the factors used cannot be changed (e.g. capital equipment)
- **variable factors of production:** resource inputs that can be varied in the short run (e.g. raw materials), when at least one factor of production is fixed
- **production function:** the ratio of inputs to output over a given time period; it shows the resources needed to produce a maximum level of output, assuming that the inputs are used efficiently
- **total product:** the total output produced by the factors of production
- **average product:** the output per unit of the variable factor (e.g. output per worker per period of time); also referred to as 'productivity'
- **marginal product:** the additional output that is produced from employing another unit of a variable factor (e.g. the extra output from employing an additional worker)

In the short-run period of the production process, there will be at least one **fixed factor of production** or resource input. For example, it will be difficult to change the number of machines in a factory in the short run.

It will be possible, however, to change the quantity of **variable factors of production** in the short run. For example, it will be relatively easy to buy in more component parts and raw materials in the short run.

The **production function** shows the relationship between inputs and output over a given time period. It indicates how a given level of output is produced as a result of using the various factors of production involved in the production process.

🔗 It is difficult to be precise about just how long a period of time the 'short run' is; it is simply defined as a period of time when at least one factor of production is fixed.

🔗 **Time:** it is important to distinguish clearly between the short run and the long run-in relation to the production process. The short run is the period of time when at least one factor of production is fixed, whereas the long run is the period of time when all factors of production are variable.

The definition and calculation of total product, average product and marginal product

It is important to distinguish clearly between these three different types of product:

» **Total product** is the total output resulting from the use of the factors in the production process over a period of time; it is calculated by adding together the total output produced by the factors of production.
» **Average product** is the output per unit of the variable factor, such as the output per worker per period of time. This is also known as 'productivity'. It is calculated by dividing the total output by the variable factor involved to give a per unit figure.
» **Marginal product** refers to the additional or extra output that is produced as a result of employing one more variable factor (e.g. one more worker). It is calculated by dividing the extra output from employing another unit of a variable factor by that variable factor.

🔗 Questions on the different types of product can often feature in multiple-choice questions in Papers 1 and 3, so it is important that you are able to distinguish clearly between them. For example, the distinction between average and marginal data is very important.

🔗 **The margin and decision making:** the concept of the margin is important in terms of distinguishing between marginal product, total product and average product. These concepts are important in relation to how a firm makes its production decisions. Whereas *total product* refers to the total output produced by a firm, *average product* refers to the output per variable factor (i.e. the productivity of that factor), while *marginal product* refers to the extra output from employing another unit of a variable factor, such as employing one additional worker.

The law of diminishing returns or law of variable proportions

- **law of diminishing returns:** as additional units of a variable factor (e.g. labour) are added to a fixed factor (e.g. capital), the additional output (or marginal product) of the variable factor will eventually diminish; also known as the 'law of variable proportions'

» In the short run, the production process involves the combination of fixed and variable factors. Additional units of a variable factor, such as labour, could be employed and combined with a fixed factor of production, such as capital equipment.
» The **law of diminishing returns** states that, although the total output (or total product) is still increasing, it will increase at a diminishing rate. This is because the extra output, resulting from the employment of the additional worker, will eventually diminish.
» The law of diminishing returns is also known as the 'law of variable proportions'.

🔗 It is important to remember that the law of diminishing returns is a concept used in analysing the short run, when one factor of production (e.g. capital) remains fixed.

The short-run cost function

The definition and calculation of fixed costs and variable costs

fixed costs: the costs of production that remain constant at all levels of output, including zero production (e.g. rent and interest payments)
variable costs: the costs of production that vary with changes in output; the cost is zero if nothing is produced (e.g. the cost of raw materials and component parts)

» It has already been pointed out that it is important to be able to distinguish clearly between fixed and variable factors of production. It is also important to distinguish between **fixed** and **variable costs** of production.
» Variable costs are costs that vary with changes in output. If output is zero, there will be no need to pay for raw materials or component parts, so these are examples of variable costs.
» On the other hand, even if output is zero, there will be some costs of production, such as the cost of renting a factory or the interest payments on a loan. These are fixed costs, which remain constant at all levels of output in the short run.

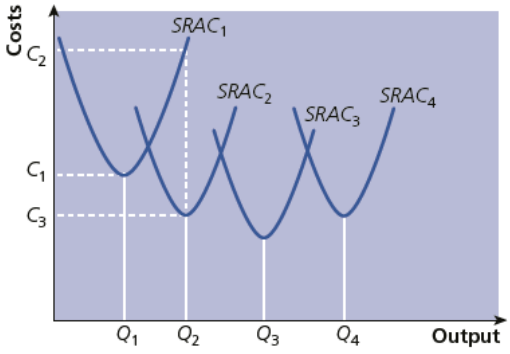
The definition and calculation of total, average and marginal costs

total cost: the sum of all costs incurred by a firm in producing a particular level of output
average cost: the total cost of employing all the factor inputs divided by the number of units produced; also known as 'average total cost'
marginal cost: the additional cost of producing an extra unit of a product
average fixed cost: the total fixed cost of production divided by the number of units produced
average variable cost: the total variable cost of production divided by the number of units produced

It is important to distinguish between these three different types of cost:

» **Total cost** is the sum of all costs resulting from the use of the factors in the production process.
» **Average cost** is the total cost of production divided by the number of units that are being produced.
» **Marginal cost** is the addition to the total cost of producing one additional unit and can be calculated by dividing the change in total cost by the change in output.

The short-run average cost curve



F7.8 Short-run cost curves with different levels of capital input

» Figure 7.8 shows a number of short-run average cost (*SRAC*) curves. Each of these is U-shaped and shows the relationship between output and cost on the basis that capital is fixed in the short run.
» Each is U-shaped as a result of decreasing **average fixed cost** and increasing **average variable cost** as output is increased.
» If a firm wishes to change output, it will have to change the amount of the variable factor being used, such as labour.
» The position of the average cost curves in the short run will therefore depend on the quantity of capital being used in the production process — in other words, there is a short-run average cost curve for each quantity of capital used, such as *SRAC*1 and *SRAC*4.

The marginal cost curve

The marginal cost curve is U-shaped. This is because when a firm initially increases its output, total costs, as well as variable costs, start to increase at a diminishing rate. At this stage, marginal cost falls until it is at its minimum. Then, as output continues to rise, marginal cost increases.

🔗 It is important to understand that the marginal cost curve will always cross the average cost curve at its lowest point. The reason for this is that when marginal cost is less than average cost, average cost will be falling; when marginal cost is more than average cost, average cost will be rising.

🔗 **The margin and decision making:** another example of the importance of the concept of the margin can be seen in relation to the existence of marginal cost in the production process. The concept of marginal cost is important in the theory of the firm because it helps the firm make its production decisions. The marginal cost curve, above where it crosses the average variable cost curve, is the firm's supply curve in the short run.

The long-run production function

- **returns to scale:** the relationship between the level of output produced by a firm and the quantity of inputs required to produce that output

There are two distinctive characteristics of the long-run production function:

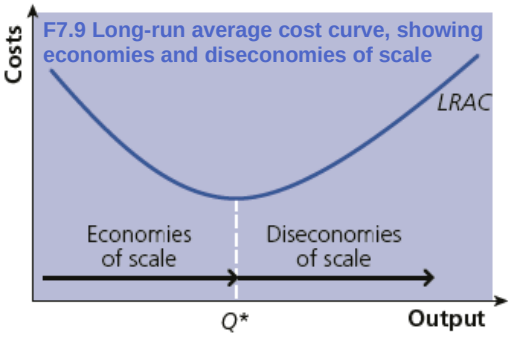
» It takes into account the fact that all factors of production are variable in the long run — that is, there are no fixed factors of production — unlike the situation with the short-run production function.
» The relationship between a firm's level of output and the quantity of units of factors of production (factor inputs) needed to produce it can change in the long run. A business can change its scale of production and, if there are:

increasing **returns to scale**, then economies of scale are being experienced.

The long-run cost function

The explanation of the shape of the long-run average cost curve

Just as the short-run average cost curve (*SRAC*) is U-shaped, it is the same with the long-run average cost (*LRAC*) curve. This can be seen in Figure 7.9. As output increases, there is initially a decrease in the average cost of production and this is shown by the falling part of the *LRAC* curve. The curve reaches its minimum point at *Q** and then starts to rise again, indicating an increase in the average costs of production.



F7.9 Long-run average cost curve, showing economies and diseconomies of scale

The concept of the minimum efficient scale

- **minimum efficient scale (MES):** the lowest level of output where average cost is at the minimum

The level of output where the *LRAC* is first at its lowest point is known as the **minimum efficient scale (MES)** —in other words, achieving MES minimises long-run average total cost. It is the minimum quantity of output at which internal economies of scale are fully exploited; no further economies of scale can be achieved beyond this scale of operation.

The relationship between economies of scale and decreasing average costs

- **economies of scale:** the benefits gained from a fall in long-run average costs of production as the scale of operations grows and the output of a firm increases

In Figure 7.9, the *LRAC* curve is falling up to a point when it is at its minimum. These decreasing average costs of production resulting from an increase in output are known as **economies of scale**. This process is also referred to as increasing returns to scale.

It is possible to distinguish between internal and external economies of scale.

Internal and external economies of scale

Internal economies of scale

- **internal economies of scale:** the advantages of a firm growing in size in the form of a reduction in the average cost of production
- **financial economies:** a reduction in average cost as a result of a larger firm being able, for example, to negotiate more favourable borrowing terms on a loan
- **technical economies:** a reduction in average cost as a result of the application of advanced technology in a firm, which brings about a greater degree of efficiency
- **economies of large dimensions:** a reduction in average cost as a result of using larger factors of production (e.g. larger containers in the transportation process)
- **risk-bearing economies:** by diversifying into different markets, the overall pattern of demand is more predictable, so a firm can save on costs (e.g. by reducing the amount of stocks held in reserve)
- **diversification:** where a firm decides to operate in a number of markets to spread risk

Decreasing average costs of production can come about as a result of a firm becoming larger. There are various examples of these **internal economies of scale:**

Type of internal economy of scale	Explanation
Financial	Larger firms may be able to benefit from negotiating a loan at a lower rate of interest than a smaller firm, mainly because financial institutions regard them as less of a risk.
Purchasing	Larger firms are often able to take advantage of their size by negotiating favourable terms with suppliers as a result of bulk buying.
Managerial	As a firm increases in size, its employment of managers does not usually need to grow at the same rate, so the cost of management per unit of output is likely to fall. Larger firms can also employ specialist managers and this is likely to increase the efficiency of a firm.
Technical	Economies of increased or large dimensions can come about more easily in a large firm, reducing costs of production — for example, if a firm doubles the size of containers that it uses, it is not likely to lead to a doubling of the cost.
Marketing	A larger firm may be better able to negotiate a lower rate — for example, in relation to buying advertising time on television.
Risk bearing	A larger firm is likely to be more diversified, spreading risks across a wider range of markets. This enables average costs to be reduced. For example, **diversification** will lead to more stable and predictable demand, and this means that a firm will not need to keep as much stock as before, reducing the costs of stockholding.

KEY SKILL

Analysis: you need to be able to analyse the different ways in which a firm can benefit from the existence of internal economies of scale — for example, the ability of a firm to benefit from a higher return on capital investment and the possibility of an increase in profitability.

External economies of scale

- **external economies of scale:** when costs of production fall because of developments outside a particular firm

The fall in the average costs of production of a firm can also be the result of external, rather than purely internal, factors.

Examples of **external economies of scale:**

Type of external economy of scale	Explanation

Concentration	If a number of firms are located in a particular area, this may encourage the development of specialist ancillary or support firms in the area that supply component parts to the firms, reducing the average costs of production.
Specialised labour	A pool of specialised skilled labour may be available in a particular area, which all firms in that industry in the area can benefit from.
Knowledge	Firms in an industry may benefit from specialist research or marketing agencies, and so be able to reduce the costs of acquiring and using this information.

🔗 Candidates need to demonstrate they clearly understand the distinction between *internal economies of scale*, which relate to a particular firm, and *external economies of scale*, which relate to a much wider aspect, such as a whole industry.

KEY SKILL

Diagrams: whereas internal economies of scale involve a movement down along a long-run average cost curve, external economies of scale will involve a downward shift of the whole long-run average cost curve.

Returns to scale

- **constant returns to scale:** where average costs remain the same as the level of output increases
- **increasing returns to scale:** where an increase in factors of production leads to a more than proportionate increase in output
- **decreasing returns to scale:** where an increase in factors of production leads to a less than proportionate increase in output

Constant returns to scale indicate a situation where average costs remain the same even as the output produced increases. In this situation, a firm will add factors of production in the same proportion so that there are constant additions to total output.

Increasing returns to scale indicate a situation in which output can be increased using a proportionately smaller quantity of units (i.e. it refers to a change in output, not cost).

Decreasing returns to scale indicate a situation in which an increase in all factors of production leads to a less than proportionate increase in output.

Internal and external diseconomies of scale

Diseconomies of scale

diseconomies of scale: where the same proportional increase in productive factors gives rise to decreasing additions to total output; also known as ‘decreasing returns to scale’

the long-run average cost curve, after initially falling, begins to rise as output is increased beyond Q^*. This indicates that there are **diseconomies of scale**. As with economies of scale, it is possible to distinguish between internal and external diseconomies of scale.

Internal diseconomies of scale

- **internal diseconomies of scale:** the disadvantages of a firm growing in size, resulting in an increase in the average cost of production

Increasing average costs of production can come about as a firm grows too large.
There are various examples of these **internal diseconomies of scale:**

Type of internal diseconomy of scale	Explanation
Problems with communication	There may be poorer communication in a large firm, which adversely affects efficiency.
Problems with motivation	There may be lower levels of motivation in a large firm, with some employees feeling alienated. This may lead to a greater likelihood of industrial disputes.
Problems with management	Problems of poor communication and low levels of motivation can lead to difficulties in managing a firm effectively.
Problems with flexibility	Larger firms may become less flexible, making it more difficult to respond to changing market conditions.

External diseconomies of scale

- **external diseconomies of scale:** when average costs of production rise because of the growth of an industry

The rise in the average costs of production of a firm can also be the result of external, rather than purely internal, factors.

Examples of **external diseconomies of scale:**

Type of external diseconomy of scale	Explanation
Competition for inputs	The increase in the size of firms, and indeed the whole industry, can lead to greater competition for a limited number of productive inputs and, as a result, the cost of these inputs may increase. For example, there could be an increase in the cost of labour or land.
Congestion	If a number of firms tend to be located in a particular area, this can lead to greater congestion, reducing the efficiency and effectiveness of the transport system and leading to an increase in costs.

The definition and calculation of revenue: total, average and marginal revenue

- **total revenue (TR):** the total amount of income received from sales of a product, calculated as the number of units sold multiplied by the price of each unit
- **average revenue (AR):** the total revenue obtained by a firm from sales divided by the number of units sold

- **marginal revenue (MR):** the extra revenue obtained by a firm from the sale of an additional unit of a product

It is important to distinguish clearly between these three different types of revenue:

» **Total revenue (*TR*)** refers to the money received from sales of a product. It is calculated using the equation:

$$TR = Q \times P$$

» **Average revenue (*AR*)** indicates the total revenue obtained from selling a product divided by the number of units sold. It is calculated using the equation:

$$AR = \frac{TR}{Q}$$

» **Marginal revenue (*MR*)** refers to the additional revenue received when one more unit of a product is sold. It is calculated using the equation:

$$MR = \frac{\text{change in TR}}{\text{change in Q}}$$

🔗 **The margin and decision making:** the concept of the margin is important in the production process in relation to marginal revenue. You need to be able to distinguish clearly between marginal revenue, average revenue and total revenue.

The definition of normal, subnormal and supernormal profit

- **profit maximisation:** the situation where marginal cost is equal to marginal revenue
- **profit:** the reward to enterprise, defined as the difference between total revenue and total costs
- **break-even point:** the level of output at which a firm is making neither a loss nor a profit

It is generally assumed that the main objective of a firm is **profit maximisation**.

Profit is defined as the difference between the total revenue received by a firm and the total costs involved in producing what is sold — in other words, it is equal to total revenue minus total costs. Profit maximisation is a situation where marginal cost is equal to marginal revenue.

The **break-even point** is the level of output at which a firm is making neither a loss nor a profit.

It is important, however, to distinguish between normal profit, subnormal profit and supernormal profit.

🔗 **The margin and decision making:** the concept of the margin is very significant in terms of profit maximisation because this refers to the situation when marginal cost is equal to marginal revenue.

Normal profit

Normal profit refers to the amount of profit that needs to be made by a firm to stay in a particular market. It is the profit necessary to just cover the opportunity cost of the resources used in production. It is actually included in the average cost curve of a firm. It is where price is equal to average cost ($AR = AC$) and where total revenue equals total cost.

Subnormal profit

Subnormal profit is any profit less than normal profit and where $AR < AC$. It is where average cost is greater than price. In the short run, a firm will continue producing while making subnormal

profit (i.e. making a loss), as long as the average variable cost (*AVC*) is covered. However, in the long run, the firm will close down.

Supernormal profit

- **normal profit:** the level of profit that a firm requires to keep operating in the industry
- **subnormal profit:** any profit less than normal profit
- **supernormal profit:** the level of profit over and above normal profit; also known as 'abnormal profit'

A firm may, however, wish to make a profit that is above normal profit. This level of profit that is over and above normal profit is known as **supernormal profit**. It is where price is greater than average cost (*AR* > *AC*).

🔗 Candidates need to be able to distinguish clearly between *normal profit*, *subnormal profit* and *supernormal profit*, both in terms of how they are shown in diagrams and how they can be defined using *AR* and *AC*.

The economist's and the accountant's definition of profit

As an extension point, it is useful to understand that an economist's definition of profit will be different from that of an accountant.

» The economist's definition of normal profit includes within it a rate of return that is needed for the firm to remain in that particular market or industry.
» An accountant, however, takes a different view, pointing out that this assumed rate of return cannot be formally and explicitly identified in the accounts of a firm. An accountant is concerned with the actual figures that can be produced by a firm as a resul t of its various trading activities.
» An economist is prepared to build into their assumptions that a firm will need to make a certain profit that is just enough to keep it in its present line of business.
» For an accountant, however, the fact that it is impossible to say precisely what the size of this normal profit needs to be is sufficient for such a figure not to be included in any accounting data.

🔗 Candidates need to know that an economist and an accountant will consider profit from different points of view.

The calculation of supernormal and subnormal profit

Supernormal profit

Supernormal profit occurs when total revenue is greater than total cost and is calculated by total revenue minus total cost (including both the fixed and the variable costs). The total costs include a reward to all the factors, including normal profit.

Subnormal profit

Subnormal profit occurs when total cost is greater than total revenue and is calculated by total revenue minus total cost. As total cost is greater than total revenue, this will give a negative figure.

7.6 Different market structures

Perfect competition and imperfect competition

It is important to distinguish between the different types of market structure that can occur in an economy. There are five main market structures:

» perfect competition
» monopolistic competition
» oligopoly
» monopoly
» natural monopoly

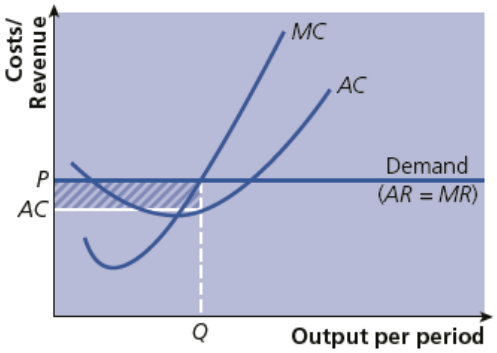
The last four structures are all examples of imperfect competition.

Perfect competition

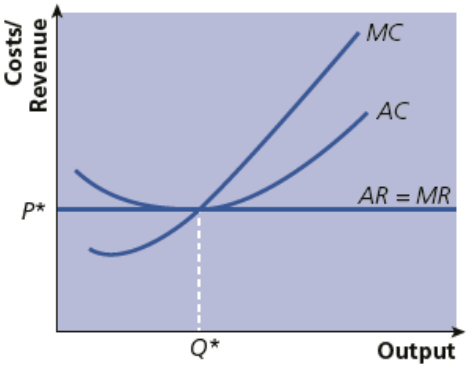
- **perfect competition:** a market or industry consisting of many virtually identical firms which all accept the market price in the industry
- **barriers to entry or exit:** various obstacles that make it very difficult, or impossible, for new firms to enter or exit an industry (e.g. technical economies of scale, patents or heavy capital investments)

The market structure of **perfect competition** is based on a number of assumptions, which include the following:

» There are many buyers and sellers.
» The buyers and sellers are price takers — that is, they just have to accept the price that prevails in the market and are unable to influence it in any way.
» There is perfect knowledge among producers and consumers, so the buyers and sellers know which products are for sale and at what price.
» The product is homogeneous, so there is no possibility of product differentiation.
» There are no **barriers to entry or exit**, so firms can enter or leave the industry in the long run, if they so wish.
» There is perfect factor mobility in the long run.
» There are no transport costs.
» All producers have access to the same technology.
» Each firm faces a perfectly elastic demand curve for its product (although the demand curve for the whole industry is downward sloping from left to right).
» Firms aim to maximise profits.
» Only normal profit can be earned in the long run.



F7.10 A firm making supernormal profit in the short run under perfect competition



F7.11 Long-run equilibrium under perfect competition

In the short run, a firm could make supernormal profits, normal profits or subnormal profits. If it were unable to cover its short-run average variable cost, it would need to exit from the market.

Figure 7.10 shows the equilibrium situation of a firm in perfect competition that is making supernormal profits. The profit maximisation position, where *MC* equals *MR*, is at output *Q*, but at this point the price is shown by *P* and the average cost by *AC*. At this point, *AR* is greater than *AC* and this explains the existence of supernormal profits in the short run, shown by the shaded area.

In the long run, if a firm makes supernormal profits, new firms will be attracted into the industry. If, however, a firm makes subnormal profits (that is, where *AC* is greater than *AR*), it will decide to leave the industry. As a result of such changes, only normal profits will be made by the firms in the industry in the long run.

Figure 7.11 shows the equilibrium situation of firms in perfect competition in the long run. *AR* is equal to *AC* at a price of *P** and a quantity of *Q** and so only normal profits are made.

Perfect competition is very much a theoretical model of the behaviour of firms, based on a number of assumptions. It has therefore been criticised for being unrealistic, but it is still useful as a model of economic behaviour. There are some markets that come relatively close to the features of perfect competition, such as agricultural markets and street food vendors.

🔗 It would be relatively easy for candidates to dismiss the theory of perfect competition for being unrealistic, but the theory is important in providing a benchmark to consider other theories of market structures.

Imperfect competition

- **imperfect competition:** a market that lacks some, or all, of the features of perfect competition

imperfect competition covers all of the following:

» monopolistic competition
» oligopoly
» monopoly
» natural monopoly

Monopolistic competition

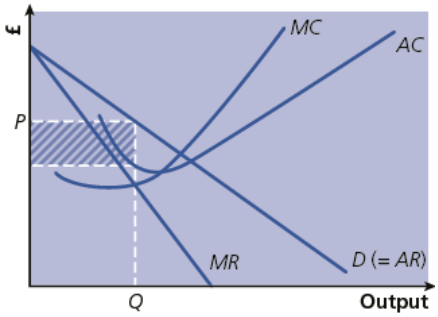
- **Monopolistic competition:** a market or industry where there is competition between a large number of firms that produce products that are similar but differentiated, usually through the use of brand images

The market structure of **monopolistic competition** is based on a number of assumptions, which include the following:

» There are a large number of firms (although fewer than in perfect competition).
» Products are differentiated — they are not homogeneous.
» Each firm faces a demand curve that is downward sloping from left to right.
» Demand for a product is relatively price elastic, but not perfectly elastic.
» Great use is made of advertising brand images to build up brand loyalty.
» There are no barriers to entry and exit, so firms can enter or leave the industry in the long run, if they so wish.
» Firms aim to maximise profits.
» Only normal profits can be earned in the long run.

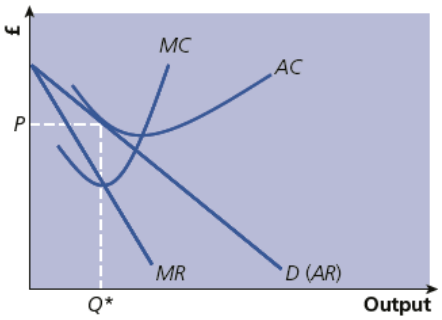
In the short run, a firm could make supernormal profits, normal profits or subnormal profits. As there are no barriers to entry or exit, this will enable firms to leave or enter the industry in response to these profits.

F7.12 shows the equilibrium situation of a firm in monopolistic competition that is making supernormal profits. The profit maximisation position, where *MC* equals *MR*, is at output *Q* and price *P*. *AR* is above *AC*, so the firm is making supernormal profits, shown by the shaded area.



F7.12 Short-run equilibrium under monopolistic competition

In the long run, if a firm makes supernormal profits, new firms will be attracted into the industry. If, however, a firm makes subnormal profits — that is, where *AC* is greater than *AR* — it will decide to leave the industry. As a result of such changes, only normal profits will be made by the firms in the industry in the long run.



7.13 Long-run equilibrium under monopolistic competition

Figure 7.13 shows the equilibrium situation of firms in monopolistic competition in the long run. *AR* (or *D*) is equal to *AC* at a price of *P* and a quantity of *Q** and so only normal profits are made.

Monopolistic competition is generally regarded as a more realistic model of the behaviour of firms than perfect competition because it is often the case that there is a great deal of competition between firms that are selling products that are similar, but not identical

🔗 in some questions which require candidates to compare monopolistic competition and monopoly, it is not always clear when a candidate is referring to monopolistic competition and when they are referring to monopoly. It is important that candidates clearly indicate which market structure they are referring to throughout their answer.

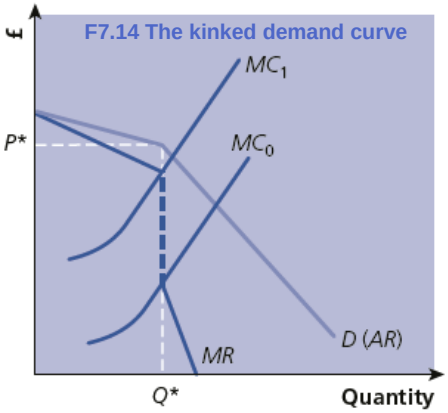
Oligopoly

- **oligopoly:** a market or industry in which there are a few large firms competing with each other

The market structure of **oligopoly** is based on a number of assumptions, which include the following:

» There are a small number of firms (if there are only two firms, it is called a *duopoly*).
» There are differentiated products.
» Great use is made of advertising brand images to build up brand loyalty.
» Barriers to entry make it difficult for new entrants to enter the market.
» Firms can make supernormal/abnormal profits in the long run, as well as in the short run.
» There could be a mixture of price makers and price takers.
» The firms are interdependent.
» There could be a degree of collusion between firms operating in a cartel.
» There is a kinked demand curve.
» There is a great deal of price stability/rigidity.

An important distinctive feature of oligopoly is the existence of a kinked demand curve. This results from the fact that firms in an oligopolistic market structure try to anticipate the reactions of rival firms to their actions.



F7.14 The kinked demand curve

The kink in the demand curve can be explained as follows:

» **Above the kink**: it is assumed that if an oligopolistic firm increases its price, other firms in the market will not follow, so demand above the kink is elastic.
» **Below the kink**: it can also be assumed that if an oligopolistic firm reduces its price, other firms in the market will follow, so demand below the kink is inelastic.

Figure 7.14 shows the equilibrium situation of oligopolistic firms.

The profit maximisation position, where *MC* equals *MR*, is at output *Q** and price *P**. Above the kink (i.e. to the left of *Q**), demand is elastic; below the kink (i.e. to the right of *Q**), demand is inelastic. It should be noted that the marginal revenue line is discontinuous, shown by the dotted line between *MC*1 and *MC*0.

Oligopoly is generally regarded as a reasonably realistic model of the behaviour of firms, but it is more complex than the other market models because there are many ways in which firms in such a market may interact with each other.

Monopoly

- **monopoly:** a market or industry where there is a single firm which controls the supply of the product

The market structure of **monopoly** is based on a number of assumptions, which include the following:

» There is one firm in an industry — that is, just one single seller.
» Legally, a monopoly can be defined in terms of a percentage of market share — for example, in the UK it is 25%.
» The monopoly is a price maker.
» There are no substitutes for the product.
» Barriers to entry make it virtually impossible for new firms to enter the market.
» Supernormal profits can exist in both the short run and the long run because of the existence of barriers to entry, preventing them from being competed away.
» The demand for the firm's product is also the market or industry demand, so the demand curve is downward sloping from left to right.
» Marginal revenue is always less than average revenue.
» The firm aims to maximise profits.

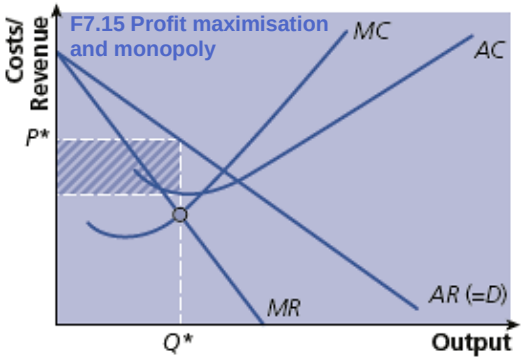


Figure 7.15 shows the equilibrium situation of a monopolistic firm. The profit maximisation position, where *MC* is equal to *MR*, is at price *P** and quantity *Q**. Supernormal profits are being made, but because of the existence of barriers to entry, it is impossible for new firms to enter the market. As a result, these profits can exist in the long run as well as the short run.

A monopoly is generally regarded as a market structure that has disadvantages for the consumer, but this need not always be the case, as in the case of a natural monopoly.

Natural monopoly

- **natural monopoly:** a situation where average cost will be lower with just one provider, avoiding the wasteful duplication of resources

A **natural monopoly** exists where a single supplier has a very significant cost advantage as a result of being in a monopoly situation. If there were competition with other producers in the market, the costs of production would actually increase. In such a situation, the monopoly can be described as natural because it avoids the extra costs that would come about as the result of a duplication of resources.

It may well be the case in a natural monopoly that one firm will have sufficient economies of scale to satisfy market demand more efficiently than two or more firms.

🔗 In examination questions on monopoly, especially those which require candidates to discuss or evaluate a monopoly market, it is important to include *both* the potential advantages *and* the possible disadvantages of monopoly before coming to a conclusion.
🔗 **Equilibrium and disequilibrium:** the different market structures can be compared in relation to positions of equilibrium and disequilibrium in the various markets.

The structure of markets

the main distinguishing features of the **four types of market structure**:

	Perfect competition	Monopolistic competition	Oligopoly	Monopoly
Number of buyers and sellers	Many buyers and sellers	Many buyers, but the number of sellers is not as large as under perfect competition	Many buyers, but few sellers	Many buyers, but one seller
Degree of freedom of entry	Not restricted	Not restricted	Some barriers to entry, especially in the long run	High barriers to entry
Firm's influence over price	None; the firm is a price taker	Some	Some	Price maker, subject to the demand curve
Product differentiation	Homogeneous	Differentiated	Differentiated	No close substitutes
Availability of information	Perfect knowledge	Lack of perfect knowledge	Lack of perfect knowledge	Lack of perfect Knowledge
Examples	Cauliflowers; carrots	Fast-food outlets; travel agents	Cars; mobile phones	PC operating systems; the market for local water supply

Barriers to entry and exit

Barriers to entry into, and exit from, markets are the various factors that can prevent firms entering or leaving a market, or make it difficult for them to do so. The existence of barriers to entry and exit makes a market less competitive; the greater the barriers that exist, the less competitive a market will be.

There are a number of barriers to entry and exit, including the following:

» **Legal barriers:** a patent can act as a legal barrier to entry into a market because other firms will not have permission to produce and sell a particular product.
» **Market barriers:** in some markets, a great deal of money is spent on establishing a strong brand image and the expense of such advertising could act as a barrier to entry.

» **Cost barriers:** economies of scale occur when increased output leads to lower average costs, so new firms, with relatively low output, will experience relatively higher costs and this can act as a barrier to entry.
» **Physical barriers:** some countries will have supplies of a particular product, but other countries will not (e.g. oil reserves are only found in certain countries), and the absence of such resources will act as a barrier to entry into this market.

The performance of firms in different market structures

- **x-inefficiency:** a situation where average cost is not at its lowest point because monopoly power has given rise to inefficiency
- **contestable market:** a situation where it may be relatively easy for new entrants to enter a market or industry; the effect of this is that existing firms in an industry face the threat of new firms coming into the industry and increasing the degree of competition
- **sunk costs:** costs which were paid when a firm entered a market and are non-recoverable when it leaves, e.g. for research and development

The performance of firms in different market structures can be compared in various ways, including the following:

» **Revenues and revenue curves**: the *AR* and *MR* curves are horizontal in perfect competition (the *AR* curve is also the firm's demand curve because *AR* is the amount of revenue per unit sold and this is equal to the price at which the product is sold), but are downward sloping in monopolistic competition, oligopoly, monopoly and natural monopoly, with the *MR* curve below the *AR* curve (in oligopoly, the *AR* curve is kinked and the *MR* curve is discontinuous).
» **Output in the short run and the long run**: the profit-maximising output in all of the market structures is determined where *MC* = *MR*.
» **Profits in the short run and the long run**: it is possible for firms in all market structures to make supernormal profits in the short run, but in perfect competition and monopolistic competition these will be competed away in the long run and only normal profits will be made; this will not be the case with the other market structures.
» Shutdown price in the short run and the long run: a firm can make subnormal profits in the short run and continue in production as long as average variable costs are being covered, otherwise it will be forced to shut down; in the long run, a firm will need to make at least a normal profit — price will need to equal average cost.
» **Derivation of a firm's supply curve in a perfectly competitive market**: in the short run, a firm's supply curve in perfect competition is its *MC* curve above the point where it intersects the *AVC* curve; in the long run, a firm's supply curve in perfect competition is its *MC* curve above the point where it intersects the *ATC* curve (i.e. equal to or above its break-even point).
» **Efficiency and x-inefficiency in the short run and the long run**: firms in perfect competition are efficient in the long run, in terms of both productive efficiency and allocative efficiency; in the other market structures, there is
x-inefficiency — that is, production is not at the minimum average cost and price is not equal to marginal cost.

🔗 **Time:** time is an important concept in relation to market structures because there can be significant differences in terms of what can happen in the short run and the long run in different markets.

🔗 **Efficiency and inefficiency:** firms in different market structures can be compared in terms of their efficiency and inefficiency.

Contestable markets

The key feature of a **contestable market** is that because it is relatively easy for a new firm to enter a market, the existing firms in the market are continually influenced by the threat of the possible entry of such firms, making them behave as if these potential firms were actually already operating in the market.
The idea of contestable markets is based on a number of assumptions, including the following:

» There are no barriers to entry or exit; entry into, and exit from, an industry is relatively easy and costless.
» Firms already in the market continually face the threat of competition.
» This puts pressure on the firms to be efficient.
» Supernormal profits are made in the short run, but only normal profits in the long run.
» There are degrees of contestability; a perfectly contestable market will have no **sunk costs** — costs that a firm paid when it entered the market, and which are non-recoverable when it leaves, such as research and development.
» There is perfect information.
» All firms, including those in a market and those planning to enter a market, have access to the same level of technology.

There are a number of implications of contestable markets:

» The existence of low barriers to entry and exit means that new firms can come into a market to provide competition to established firms.
» This threat of market entry means that existing firms will be efficient in terms of productive, allocative and dynamic efficiency.
» The absence of sunk costs reduces the risk involved in entering a market.
» The continuous possibility of new firms entering a market causes existing firms in a market to focus more on sales maximisation than profit maximisation.
» The size and number of firms is irrelevant.

🔗 It is important that candidates clearly understand that in a contestable market, it is not so much the entry of new firms into a market that is important but the *threat* of such entry.

Also, it is important that candidates understand that a contestable market is not actually a separate market structure, but a feature that can exist to different degrees in each of the other market structures.

Price competition and non-price competition

Price competition

- **price competition:** a process of setting competitive prices to achieve particular objectives in a market

Price competition is one of the ways that a product can compete in the marketplace. It refers to a situation in which firms try to sell their products at lower prices than similar products sold by other firms. Firms develop different price strategies to beat their competitors, usually setting the same or a lower price than that charged by the competitors.

🔗 Price competition is not used in perfect competition because all of the firms in that market structure are price takers. It is also not used very often in oligopoly.

Non-price competition

- **non-price competition:** a process of using a variety of ways to increase sales other than price

The conduct of firms can be compared in relation to whether they rely on pricing policies or non-price policies. In most markets, firms compete in terms of price, but in monopolistic competition and oligopoly there is a great deal of **non-price competition** — rather than using changes in price to compete, the firms use other ways to compete with each other.
Non-price competition involves any form of competitive activity, other than changes in price, and can include:

» advertising and product promotion
» branding and the creation/maintenance of brand image and customer loyalty
» sales promotions (e.g. through such special offers as BOGOF, 'buy one, get one free')
» distribution (e.g. controlling the distribution of products to particular retail outlets or particular internet sites)
» distinctive/exclusive packaging
» differences in quality or design
» establishment of a unique selling point (USP) through product differentiation
» warranties and after-sales service
» the provision of a loyalty card which gives different types of reward
» establishment of a clear ethical and/or environmental reputation

Collusion and the prisoner's dilemma in oligopolistic markets, including a two-player pay-off matrix

- **price agreement:** where firms in an oligopolistic market agree to fix prices between themselves
- **cartel:** where a number of firms agree to collude, such as by limiting output to keep prices higher than would be the case if there were competition between them
- **prisoner's dilemma**: a situation where two prisoners must decide whether to confess, without knowing whether the other will confess or not

Collusion

It has already been stressed that a key feature of oligopoly is that firms may sometimes act together — in other words, there might be collusion between two or more firms which cooperate for their mutual benefit. Firms in oligopoly are interdependent and so will consider the likely reactions of the other firms in a market. Collusion is a way of lowering some costs so as to maintain supernormal profits, reducing consumer welfare.

This collusion can take different forms, including a price agreement and a cartel.

» A **price agreement** is where firms in an oligopolistic market agree to fix prices between themselves.
» A **cartel** is where a number of firms agree to work together, such as by limiting output to keep prices higher than would be the case if there were competition between them.

The prisoner's dilemma

» Firms operating in an oligopolistic market make decisions in the face of uncertainty about how their rivals will react to their moves.
» Game theory is a technique of analysing the behaviour of firms in such a situation.
» The **prisoner's dilemma** is a game theory idea, based on the situation of two prisoners who are being questioned over their guilt or innocence of a crime. They must decide whether to confess, without knowing whether the other will confess or not, and where a confession carries the possibility of a lighter punishment.
» It shows why two individuals might not cooperate, even if it is collectively in their best interest to do so.

The two-player pay-off matrix

» The prisoner's dilemma is similar to an economic situation where two firms exist in a market (i.e. there is a duopoly).
» Just as self-interest drives the prisoners in the prisoner's dilemma to confess, self-interest may make it difficult for an oligopoly firm to maintain a cooperative outcome.
» The pay-off matrix is a visual representation of all the possible outcomes that can occur when two people or two firms have to make a strategic decision, such as whether to cooperate with each other or to 'go it alone'.

KEY SKILLS

Analysis and application: you need to be able to analyse the contribution that a knowledge and understanding of collusion and the prisoner's dilemma can make to the study of firms, especially firms operating in oligopolistic markets. The prisoner's dilemma provides a framework for understanding how to balance cooperation and competition, and is a useful tool for strategic decision making.

The definition and calculation of the concentration ratio

- **concentration ratio:** the percentage of a market controlled by a given number of firms (e.g. the five largest firms in an industry might control 80% of the output of the industry)

The **concentration ratio** shows the percentage of a particular market that is accounted for by a certain number of firms. For example, it may well be that four or five firms in a particular industry have an 80% share of a market between them.

In an oligopolistic market structure, there are usually only a few large firms — in other words, the concentration ratio of the largest firms is usually very high.

🔗 The higher the concentration ratio of an industry, the more imperfect the market is. On the other hand, the lower the concentration ratio of an industry, the more competitive the market is.

7.7 Growth and survival of firms

Reasons for the different sizes of firms

- **firm:** a particular and distinct organisation that is owned separately from any other organisation
- **industry:** a collection of firms producing similar products

Small firms exist in many economies. Reasons for the continued survival of small firms include the following:

» The size of the market is small.
» The firm may be in a very specific niche market.
» The firm is providing customers with a service that requires personal attention.
» The firm may have only just started and so is relatively small at present.
» The owners of the firm prefer it to remain small.
» Small firms may receive specific financial support from government.
» Small firms may be more flexible in responding to changes in consumer demand.
» Small firms may be more innovative and pioneering.
» The small firm may not be able to grow because of difficulties in raising the necessary finance.

» In some industries, there has been an increase in the process of ‘outsourcing’ and ‘contracting out’, and small firms may benefit from this.
» A small firm may be more efficient — for example, labour relations and levels of motivation may be better than in a large firm.
However, despite the potential advantages of small firms, there are many reasons for the growth of firms, including the following:

» **To decrease costs:** to take advantage of possible economies of scale, leading to a decrease in the costs of production of a firm.
» **To reduce risk:** to be stronger and therefore safer from a hostile takeover or merger proposal (in the case of a firm becoming larger through external growth).
» **To increase profits:** a larger firm may able to gain greater profitability.
» **To fulfil management objectives:** a possible desire of owners and/or managers to expand.
» **To dominate a market:** to take advantage of opportunities to gain increased sales from a larger market share.

The concepts of firm and industry

It is important to distinguish between a firm and an industry:
» A **firm** is a distinct organisation that is owned separately from any other organisation; in a firm, an entrepreneur will bring together factors of production in order to produce particular products.
» An **industry** involves a number of firms which produce broadly similar products.

Firms can grow through two different ways:

» internal growth
» external growth

🔗 Candidates often confuse the terms ‘firm’ and ‘industry’, making it difficult sometimes for examiners to follow the logic of a candidate’s answer. It is important that you clearly understand the difference between ‘one firm’ and ‘a number of firms operating in an industry or market’.

The internal growth of firms

The internal growth of a firm comes about as a result of a firm increasing in size through producing and selling more products. The extent of this growth can be measured by:

» the volume of sales
» sales revenue (turnover)
» the number of employees
» market share
» the size of profits

The external growth of firms

- **merger:** where two or more firms combine together as a result of mutual agreement
- **takeover:** where two or more firms combine together as a result of some form of hostile bid by one firm for another; also known as an ‘acquisition’
- **integration:** the process whereby two or more firms come together through a takeover or merger
- **horizontal integration:** where firms at the same stage of production merge
- **vertical integration**: where a firm joins with another firm at an earlier stage of the production process (backward vertical integration) or a later stage of the production process (forward vertical integration)

- **conglomerate integration:** a merger between firms that are operating in completely different markets rather than in different stages of the same market
- **diversification:** where a firm decides to operate in a number of markets to spread risk

The external growth of a firm comes about as a result of a **merger** or **takeover** in which two or more firms combine together. This process is known as **integration**.

Methods of integration

Horizontal integration

One form of external growth is **horizontal integration**. This is where two or more firms at the same stage of the production process join together.

KEY SKILL

Application: examples of horizontal integration would include two banks or two car manufacturing companies.

Vertical integration

Another form of external growth is **vertical integration**. This is where two or more firms at different stages of the production process merge. If it involves going back to an earlier stage in the production process, it is known as ‘backward vertical integration’. If it involves going forward to a later stage in the production process, this is known as ‘forward vertical integration’.

KEY SKILL

Application: an example of backward vertical integration is a tyre producer taking over rubber plantations; an example of forward vertical integration is a car producer taking over a chain of petrol stations and garages to act as distributors of the vehicles.

Conglomerate integration

Both horizontal and vertical (whether backward or forward) integration involve mergers of firms operating in the same market.

In contrast, **conglomerate integration** is where two or more firms merge together even though they are operating in entirely different markets that do not always have any relationship with each other. Although this may appear illogical, it has the benefit of spreading risks in different markets through the process of **diversification**.

🔗 In examination questions on integration, candidates often write about horizontal and vertical integration only, making no reference at all to conglomerate integration. Candidates should consider all three types of integration in their answers. Conglomerate integration has the advantage of offering diversification, and this spreading of risks in different markets can be of fundamental importance to a large number of firms.

KEY SKILL

Application: an example of diversification resulting from conglomerate integration is the brewing company Guinness diversifying into publishing, producing the *Guinness Book of Records* among other titles.

The reasons for integration

The reasons for integration will depend on the type of integration:

» **Horizontal integration:** where firms operate in the same industry and at the same stage of production, they can benefit from economies of scale (e.g. two banks decide to merge).
» **Backward vertical integration:** where firms operate in the same industry, but at different stages of production, one firm might want to acquire a firm operating earlier in the supply chain (e.g. a retailer decides to buy a wholesaler).
» **Forward vertical integration:** where firms operate in the same industry, but at different stages of production, one firm might want to acquire a firm operating further up the supply chain (e.g. a vehicle manufacturer decides to buy a car parts distributor).
» **Conglomerate integration:** firms in different unrelated industries can benefit from the spreading of risk through diversification (e.g. the merger between the Walt Disney Company, a film production company, and the American Broadcasting Company, a television company, giving the Walt Disney company more extensive distribution of its products).

The consequences of integration

There are a number of possible consequences of integration, including the following:

» **Economies of scale:** integration enables firms to take advantage of different economies of scale, such as a cost savings associated with marketing and technology; this will enable them to keep costs and prices down.
» **Rationalisation:** this is the process of eliminating those parts of the operation of a business that are inefficient and unprofitable, and integration could help to bring this about
» **Sharing of knowledge:** integration enables knowledge to be shared and this could reduce or remove elements of asymmetric information.
» **Strength**: integration could send out a signal to other firms not to attempt a takeover bid; it could also strengthen anti-takeover defence/protective strategies.
» **Research and development:** integration may enable more funds to be allocated to research and development so that new innovative products can be produced, increasing the competitiveness and profitability of the business in the long run.

Cartels

The conditions for an effective cartel

The possible existence of cartels has already been referred to in relation to oligopoly. A cartel is a grouping of producers that work together to defend their shared interests. Cartels are created when a few large producers decide to collude — for example, to fix prices for members so that competition on price is avoided. They can also restrict output released onto the market.

There are a number of conditions for an effective cartel, including the following:

» **Barriers to entry:** a cartel is more likely to be effective when there are high barriers to entry into a market or an industry.
» **Control over both price and output:** a cartel is more likely to be effective if it can not only fix prices for members, but also restrict output.
» **Setting rules:** a cartel will need to set rules governing the behaviour of members; if members follow these rules, risks that would exist without a cartel are reduced.
» **Policing rules:** a cartel is more likely to be effective when all members can be ‘policed’ in some way to ensure that the rules are being obeyed.
» **Leading firm:** a cartel is more likely to be effective when there is a leading firm and homogeneous product, as well as a similar cost structure.

KEY SKILLS

Application: an example of a cartel is OPEC — the Organisation of the Petroleum Exporting Countries, consisting of 13 countries, established in 1960. It controls about 50% of global oil production and about 80% of the world's oil reserves.

Analysis: you need to be able to demonstrate a knowledge and understanding of these conditions in order to analyse the potential effectiveness of a cartel.

The consequences of a cartel

the positive effects for producers and the negative effects for consumers:

Positive effects for producers	Negative effects for consumers
• **Protection of shared interests:** members of a cartel work together to defend their shared interests. • **Avoidance of price competition:** cartels fix prices for their members, so that competition on the basis of price is avoided. • **Effect on revenue:** where a cartel is able to control both price and output, revenue will be substantially increased. • **Control of market:** the existence of a cartel will give members a dominant position in a market (e.g. OPEC controls over 50% of the oil-producing market).	• **Higher prices:** the members of a cartel can all raise prices together, which reduces the price elasticity of demand for any particular member and makes products more expensive for consumers. • **Lack of transparency:** members of a cartel may agree to hide prices or withhold information, such as in relation to hidden charges to consumers involved in certain transactions. • **Restricted output:** members of a cartel may agree to limit output onto a market, such as through a quota system, reducing competition. • **Carving up a market:** members of a cartel may collectively agree to break up a market into regions or territories and not compete in each other's area, reducing choice for consumers.

KEY SKILL

Evaluation: in evaluating the consequences of a cartel, you would need to contrast the potential advantages for producers with the potential disadvantages for consumers.

The principal–agent problem

- **principal–agent problem:** the problem that can occur as a result of the different objectives of an owner (i.e. the principal) and a manager (i.e. the agent)

It is important to understand the distinction between the ownership and the control of a firm and there may be a problem arising from this divorce of ownership from control. For example, the shareholders may own a firm, but they will not be in day-to- day control of it.

The principal is the owner and he or she will hire an agent (i.e. a manager) to run the firm on an everyday basis. The problem, however, is that the agent may not run the firm in exactly the way that the shareholders/owners would like. The two may have differing objectives for a firm. For

example, the principal may have the objective of profit maximisation but the agent may have the objective of salary maximisation.

KEY SKILL

Analysis: you need to be able to analyse why it is possible that a problem could arise between the principal of a firm and an agent, such as between the owner and a manager.

7.8 Differing objectives and policies of firms

The traditional profit-maximising objective of firms

Profit maximisation has traditionally been regarded as the main objective of a firm. Profit maximisation is defined as occurring at that level of output where the marginal cost of production is equal to the marginal revenue obtained by a firm from that level of production (*MC* = *MR*). The distinction between normal and supernormal profit was referred to in section 7.5 of this chapter.

Other objectives of firms

- **satisficing:** when the objective of a firm is to produce a level of profits that is satisfactory to stakeholders (e.g. shareholders and managers)
- **revenue maximisation:** when the objective of a firm is to maximise the total income, rather than the profits, of a firm
- **sales maximisation:** when the objective of a firm is to maximise the volume of products sold

It is necessary, however, to consider that a firm may have other objectives:

Model	Explanation
Profit satisficing	Where there is a distinction between ownership and control, managers may just want to deal in an appropriate manner with all of the stakeholders involved in a firm, so that all stakeholders are satisfied. In such a situation, satisfactory profits, rather than maximum profits, may be the aim.
Revenue maximisation	This is likely to be the objective where the salaries of managers are linked to revenue rather than to profits. In this situation, revenue is maximised when *MR* = 0.
Sales maximisation	In this situation, managers aim to maximise the volume of sales rather than the revenue resulting from such sales.
Growth maximisation	This refers to a situation where managers aim to increase the size of the firm because this will make it less vulnerable to a merger or takeover and it is also possible that salaries of managers will be linked to the size of a firm.
Survival	An objective of a firm, particularly in the first few years of its existence, may simply be to survive. This is especially the case in those markets where firms do not tend to survive for very long, such as take-away food retail outlets.
Strategic objective	A firm may establish its objectives within a broad strategic approach, such as in relation to ‘corporate social responsibility’ and the aim to operate without causing any ethical issues or environmental problems.

🔗 In examination questions on the objectives of a firm, candidates need to remember not to write exclusively on profit maximisation, but also to consider other possible objectives.

KEY SKILL

Analysis: when analysing the different objectives and policies of firms, you need to be able to consider other aims of a firm besides profit maximisation, such as survival, profit satisficing, sales maximisation and revenue maximisation. It is important that you understand why (i.e. under what circumstances) a firm could aim for an objective other than profit maximisation. It is also important that you understand that the objectives of a firm may change over a period of time for a variety of different reasons as economic circumstances change.

Price discrimination

- **price discrimination:** the process of charging different prices in different markets where there are differences in the price elasticities of demand

Price discrimination involves charging a different price to different groups of people for the same good. It is mainly practised by monopolies.

The conditions for effective price discrimination

The conditions for effective price discrimination include the following:

» **Different elasticities of demand:** there must be different price elasticities of demand in the different markets.
» **Different prices do not reflect different costs:** the differences in price should not be a reflection of differences in the costs of production.
» **Market segmentation:** the monopoly firm is able to keep different markets separate; this separation could involve different geographical regions, different times of the day or people of a different age or status.
» **No resale:** the firm needs to be able to prevent individuals in one market buying at a lower price in another market and reselling at the higher price.
» **Monopoly power:** the firm must have some control over the price — that is, it must be a price maker rather than a price taker.

Degrees of piece discrimination

It is possible to distinguish three degrees of price discrimination:

» **First degree:** this involves considering individual customers — the firm needs to be able to find out what each customer is willing to pay for a product and then sell it to the customer at that price. This means that each unit of a product is sold at a different price — for example, where a customer orders a specific item of designer jewellery and the price is negotiated between the buyer and the supplier.
» **Second degree:** this involves allowing customers to choose a particular deal — the firm offers a special deal to those customers who meet certain conditions, such as a special price for bulk purchases of a product.
» **Third degree:** this involves offering special discounts to members of certain groups, such as students, senior citizens or people travelling on trains at certain times of the day (generally known as off-peak travel).

KEY SKILL

Analysis: you need to be able to distinguish between the three degrees of price discrimination in any analysis of price discrimination.

The consequences of price discrimination

There are a number of consequences of price discrimination for both the producer and the consumer.

Consequences for the producer include the following:

» **Increase in revenue and profit:** a monopoly firm is able to increase its revenue and profit by practising price discrimination; it is able to extract consumer surplus and turn it into supernormal profit.
» **Cross-subsidisation:** a firm practising price discrimination will be able to use the supernormal profit to cross-subsidise loss-making activities in other operations that could have important social benefits.
» **Economies of scale:** an increase in total output resulting from selling extra units of a product at a lower price might help a monopoly firm to exploit economies of scale, resulting in lower long-run average costs.
» **Capacity utilisation:** price discrimination may enable more effective management of capacity utilisation by a firm, as with the example of trains.

Consequences for the consumer include the following:

» **Consumer payments:** each customer pays the price that he or she is willing to pay rather than forgo the product (the actual price paid will depend on the whether the price discrimination being practised is first degree, second degree or third degree).
» **Consumer surplus:** the consumer surplus is reduced in most cases, representing a loss of welfare; however, some consumers, who can now buy a product at a lower price, may benefit.
» **Contestable markets:** price discrimination might make a market more contestable, allowing cheap prices to be charged to certain customers.

🔗 A common error in examination answers on this part of the syllabus is to assume that differences in price are because of differences in cost. The key point to emphasise in answers to questions about price discrimination is that the price differences do *not* reflect differences in cost.

KEY SKILL

Evaluation: you need to be able to weigh up the various consequences of price discrimination for both the producer and the consumer before making a judgement on it.

Other pricing policies

- **limit pricing:** the selection of a price that is below the profit maximising price
- **predatory pricing:** where a market leader reduces prices in a deliberate attempt to force other firms out of a market

There are a number of other different pricing policies, including the following:

» **Limit pricing:** this refers to a situation where a price is selected below the profit-maximising price. A firm may decide to price in this way to discourage new firms from entering an industry. This will help to protect the competitive position of the firm that adopts this pricing strategy.
» **Predatory pricing:** oligopoly can sometimes give rise to predatory pricing. This is where a firm charges a price that is lower than those of competitors in a deliberate attempt to force other firms out of the industry.

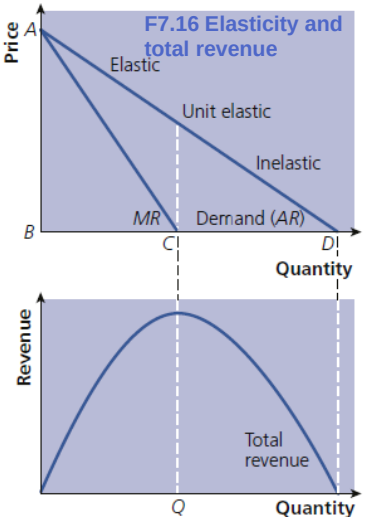
» **Price leadership:** this can happen in an oligopoly market structure where there is informal or tacit collusion between firms in the market. In this situation, firms in the market will follow the price leadership of one firm. The aim is to maximise the profits of all the firms by behaving as if they were a monopolist — that is, a single seller. As stated above, this agreement on price or output gives rise to a cartel arrangement.

🔗 It is important that you understand the difference between limit pricing and predatory pricing. Limit pricing is used to discourage new entrants into an industry whereas predatory pricing is used in relation to existing firms in an industry.

The relationship between price elasticity of demand and a firm's revenue

The relationship in a normal downward-sloping demand curve

Figure 7.16 shows the relationship between price elasticity of demand and total revenue, average revenue and marginal revenue for a downward-sloping demand curve.



If the demand curve is perfectly elastic, then marginal revenue is equal to price. If, however, there is a downward-sloping demand curve, it means that marginal revenue is less than price. This is because price has to be reduced for all products to sell just one more product. Average revenue is in fact the downward-sloping demand curve. A firm's total revenue is rising when demand is elastic, at its maximum when there is unitary elastic demand, and falling when demand is inelastic.

The relationship in a kinked demand curve

» It has already been pointed out that the demand curve in oligopoly is kinked. This occurs because firms in an oligopolistic market structure try to anticipate the reactions of rival firms to their actions.
» The kinked demand curve is an example of how the behaviour of firms can be analysed when there is no collusion between them, and it shows the mutual interdependence of firms in an oligopoly market.
» It is assumed that if an oligopolistic firm increases its price, other firms in the market will not follow and so demand above the kink is price elastic. In this situation, a fall in price will lead to a rise in total revenue.
» It is also assumed that if an oligopolistic firm reduces its price, other firms in the market will follow and so demand below the kink is price inelastic. In this situation, a rise in price will lead to a rise in total revenue.

8 Government microeconomic intervention

8.1 Government policies to achieve efficient resource allocation and correct market failure

Application and effectiveness of measures to tackle different forms of market failure

- **specific tax:** an indirect tax that is a fixed amount per unit of output
- **ad valorem tax:** an indirect tax with a percentage rate (e.g. a tax rate of 20% per product sold)

Specific and ‘ad valorem’ indirect taxes

Indirect taxes can be used to achieve the efficient allocation of resources and to correct market failure. For example, indirect taxes can be employed to discourage the consumption of demerit goods, such as tobacco.

It is important to understand that indirect taxes can be of two types:

» **Specific tax:** where a specific amount of tax is required to be paid.
» **Ad valorem tax:** where the tax to be paid is a percentage of the selling price.

Subsidies

Subsidies can be used to achieve efficient resource allocation and to correct market failure. For example, subsidies can be used to encourage the consumption of merit goods, such as education and healthcare. Chapter 3, section 3.2 covered the impact and incidence of subsidies.

Price controls

Another policy that can be used to correct market failure is the use of price controls. These are of two types:

» **Maximum price control:** if the equilibrium price of a product in a market is too high for many people to afford, a maximum price control can be established by a government to prevent the price rising above a certain level.
» **Minimum price control:** if the equilibrium price of a product in a market is too low, encouraging overconsumption, a minimum price control can be established by a government to prevent the price falling below a certain level.

KEY SKILLS

Application: an example of a product that could have a maximum price control is an essential food item, such as bread or rice. You will need to be able to analyse maximum price controls in terms of how they can improve the efficiency of resource allocation in an economy.

An example of a product that could have a minimum price control is a demerit good, such as cigarettes. You will also need to be able to analyse minimum price controls in terms of how they can improve the efficiency of resource allocation.

Problem solving: if a government decides that there is a problem with excessive consumption of demerit goods, such as alcohol or cigarettes, then it could impose a minimum price control to discourage the consumption of such products.

Application: examples of an ad valorem tax include value added tax (VAT) and goods and services tax (GST). You will need to be able to assess the effectiveness of indirect taxes, in terms of both specific taxes and ad valorem taxes, in tackling different forms of market failure.

Application: you will need to be able to assess the effectiveness of subsidies in tackling market failure, especially in terms of the underconsumption and underproduction of education and healthcare.

Production quota

- **production quota:** a limit to the quantity produced of a product over a certain period of time

» A government could decide to set a limit to the quantity of a product that may be produced in a specified time period. This is called a **production quota**.
» A quota is often used as an import control, but it could be used in a domestic economy to limit production.
» For example, if the production of a certain product is too high, lowering the market price, a quota can be used as a 'cap' on a certain level of production. If a producer exceeds this quota, a levy could be imposed on them which they would be required to pay.
» This might be the case in relation to the existence of negative externalities.

🔗 Candidates sometimes confuse a production quota and an import quota in their examination answers. A production quota involves a restriction on the domestic production of a product, whereas an import quota involves a restriction on a product entering a country from another country.

🔗 **Time:** a production quota is likely to take a certain length of time to have an impact. This time dimension is likely to reduce, or at least delay, its effectiveness.

KEY SKILL

Application: you will need to be able to analyse production quotas in terms of how they can improve the efficiency of resource allocation in an economy. For example, they could be used to reduce the extent of a negative externality, such as pollution.

Prohibitions and licences

- **prohibition:** the banning of a certain product in an economy
- **licence:** where permission to produce or sell is given by a government to a supplier, but this permission is restricted in some way

Another approach to the correction of market failure is the use of prohibitions and licences in an economy.

» A **prohibition** refers to a ban on certain products being supplied in an economy — a product subject to prohibition would be made illegal and consumers would be banned from the consumption of it. However, there might be a time lag before such a prohibition was actually put into effect.
» A **licence** is where suppliers are given permission by a government to produce or sell a product. The use of a licence gives the government some degree of control because it can limit the number of licences that it issues. However, its effectiveness may be limited by a government decision to issue only a certain number of licences.

Regulation and deregulation

Regulation

- **regulations:** a variety of legal and other rules that apply to firms in different circumstances

One way that a government could intervene to achieve an efficient resource allocation is through the use of **regulations**. A regulation refers to a law or rule that can be used by a government to reduce the extent of market failure in an economy.

There are often examples of such regulations in different economies:

» **Control of monopolies:** in some countries, a government may establish regulations to control monopolies. If a firm has too much monopoly power in a market, it can be referred to a commission (a body that can look into a monopoly situation). This can be effective by investigating whether the monopoly is acting against the public interest. Proposed mergers or takeovers might also be referred to such a regulatory body whenever it is thought that they might be against the public interest in limiting the degree of choice for consumers.

» **Consumer protection:** regulations could also exist in relation to consumer protection. In many countries, regulations have been passed that give certain rights to consumers. For example, one regulation could cover the description of a product that is being sold to ensure that consumers are properly informed. A regulatory body usually exists to ensure that such regulations are adhered to, thereby increasing their effectiveness.

» **Protection of the environment:** in some countries, regulations could be used to protect the environment, such as through controls on the level of pollution. Regulatory bodies could be given the power to fine those who are responsible for the pollution of the environment and this would certainly help to enhance their effectiveness.

Deregulation

- **deregulation:** a reduction in the number of regulations, rules and laws that operate in an industry or economy

Deregulation involves a reduction in the number of regulations that exist in an economy. The aim is to:

» allow a greater degree of competition to exist in a market than would otherwise be the case
» increase the level of efficiency as a result of the greater competition
» reduce the extent of market failure

The direct provision of goods and services

Another way in which a government could attempt to achieve efficient resource allocation and correct market failure is through the direct provision of goods and services. A government could provide goods and services itself, alongside the private sector. This is likely to be the case with certain merit goods, such as healthcare and education. It is likely to be effective as a government can use the funds it has received from taxation. In many countries, such services are provided through both the public and private sectors.

Pollution permits

- **pollution permit:** a particular type of licence that is given to a firm with the intention of reducing the level of pollution created over a period of time

» A **pollution permit** or ‘tradable permit’ is a particular example of a licence that can be issued by a government.
» The permit allows a firm to pollute the environment in some way, but only up to a certain level. This level will be less than the pollution that is taking place when the permit was first issued.
» The aim is that over time a number of permits will be issued, each allowing a lower level of pollution than the previous one. It may not be possible to eliminate the pollution completely, but issuing pollution permits should have some effect on bringing the level of pollution down over a period of time.

🔗 **Time:** the level of pollution in an economy could be reduced over a period of time by a government issuing pollution permits, each allowing a lower level of pollution than the previous one.

🔗 **Efficiency and inefficiency:** a government aims to encourage the extent of competition in a market to increase the level of efficiency.

🔗 **The role of government and the issues of equality and equity:** a government could decide to provide certain goods and services directly in an economy in order to bring about greater equality and equity. For example, it could provide certain merit goods, such as education and healthcare. This is likely to be effective because a well-educated and healthy person is likely to be able to earn a higher income than someone less educated and less healthy.

Property rights

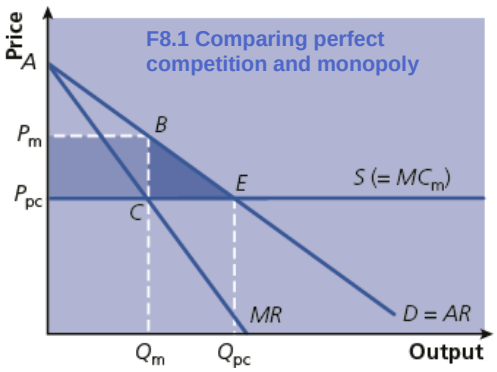
- **property right:** the right of the owner of a good to decide how it should be used

A **property right** is the right of an owner of an economic good to decide how such a good is to be used. Market failure can occur in an economy because of the absence of clear property rights. In such a situation, a government could decide to extend property rights through the encouragement of a voluntary agreement. For example, people in a community might object to the level of pollution caused by a firm in an area and an informal agreement could be reached between the community and the firm. However, if such a voluntary agreement were not successful, a government could, for example, introduce a system of pollution permits to deal with the market failure.

Nationalisation and privatisation

Nationalisation

- **nationalisation:** where a government decides to take over the ownership of a particular firm or industry
- **x-inefficiency:** a situation where average cost is not at its lowest point because monopoly power has given rise to inefficiency
- **deadweight loss:** the loss of economic efficiency that occurs when the socially optimal quantity of a product is not produced



» A government could decide to provide a particular good or service itself by providing it through a state-owned or nationalised industry.
» **Nationalisation** refers to the process by which a government takes a firm or an industry into the public sector. In other words, ownership of the firm or industry is transferred from the private sector to the public sector.
» This is often done if it is thought that such an action would be in the public interest.
» However, a nationalised industry may not always be as effective as one operating in the private sector because nationalisation involves the creation of a monopoly. This is a situation where there is just one firm in an industry and this firm can control the supply of a product in the market. it is important to see how *monopoly is an example of market failure*.
» Monopoly power is regarded as a market imperfection because the equilibrium price is likely to be higher and the equilibrium quantity lower than would be the case in perfect competition.
» Unlike the situation in perfect competition, supernormal profits are not competed away in the long run because there are significant barriers to entry, which make it very difficult, if not impossible, for new firms to enter the market.

» A particular form of inefficiency that can exist in monopoly is when a firm's average cost curve is not at its minimum level; monopoly power can mean that the lack of competition reduces the level of efficiency, so that unit costs of production are not minimised.
» This form of inefficiency is known as **x-inefficiency**. A particular example of such a market imperfection is the existence of deadweight losses.
» A **deadweight loss** arising from positive and negative externalities was covered in Chapter 7, section 7.4.
» Deadweight loss also occurs when there is a monopoly situation leading to a higher price and a lower quantity compared to the situation that would exist in a perfectly competitive market. In this sense, a consumer is worse off and so there is said to be a welfare loss.
» This can be seen in Figure 8.1, which compares the situation of perfect competition with that of monopoly. In perfect competition, the equilibrium price will be *P*pc and the equilibrium quantity will be *Q*pc. The consumer surplus will be *AP*pc*E*.
» If, however, the market becomes one with just one firm (i.e. a monopoly), the equilibrium price will now be at *P*m and the equilibrium quantity will be at *Q*m.
» The price will therefore be higher and the quantity lower than was the case in perfect competition.
» The consumer surplus is now much less, shown by the triangle *AP*m*B*.
» This creates a loss to society shown by the triangle *BCE*; this is the deadweight loss.

KEY SKILLS

Application: an example of a situation where there is an absence of clear property rights is a large open space where there are common community rights rather than private property rights.

Evaluation: you need to be able to compare the advantages and disadvantages of nationalisation before coming to a judgement as to whether it is likely to be an effective policy option or not.

Problem solving: you need to be able to consider the advantages of a firm or industry being nationalised by a government to solve a problem. For example, in some countries, loss-making private railway companies have been taken over by a government and the train service has then been provided by the state through the public sector.

Evaluation: you need to be able to weigh up the various advantages and disadvantages of privatisation before coming to a judgement as to whether it is a correct policy option to take.

Privatisation

- **privatisation:** the transfer of the ownership of an industry from the state or public sector to the private sector
- **deregulation:** a reduction in the number of regulations, rules and laws that operate in an industry or economy
- **contracting out:** the transfer of responsibility for the provision of a service from the public to the private sector
- **supply-side economics:** the approach to change in an economy that puts the focus on the supply side, rather than the demand side (e.g. privatisation, deregulation and contracting out)
- **information failure:** where people lack the full information that would allow them to make the best decisions about consumption

It has been shown that a nationalised industry may not always be as efficient as one operating in the private sector because a public sector monopoly may lead to inefficiency and the sub-optimal use of resources. A government may therefore decide to privatise an industry by transferring ownership from the public to the private sector.

potential advantages and disadvantages of privatisation:

The advantages of privatisation	The disadvantages of privatisation
• **Improvement in efficiency:** a private sector firm is likely to be more efficient and therefore more profitable. • **Greater competition:** privatisation usually leads to the creation of a number of firms in an industry, creating a greater degree of competition. • **Reduction in political interference:** decisions should be taken on the basis of economics rather than politics.	• **Creation of a private monopoly:** privatisation may not necessarily lead to greater competition and it could lead to a public sector monopoly being replaced by a private sector monopoly; this could be more difficult to regulate than a public sector monopoly. • **Focus on profits:** there may be a focus on profit rather than the public interest in the provision of essential services (e.g. healthcare). • **Fragmentation of an industry:** competition between privatised firms may not necessarily improve an industry as a whole (e.g. where a train service is provided by many different firms).

The term 'privatisation' can also apply to a number of other government initiatives, including deregulation and contracting out.

» **Deregulation** refers to the process of reducing the regulations, laws and rules which apply in an industry. When these are removed, it usually allows for a greater degree of competition to take place, which, it is argued, should lead to a greater degree of efficiency in the industry.
» **Contracting out** is another form of privatisation where a service that was originally provided by an enterprise in the public or state sector is now provided by a firm in the private sector. This, again, should lead to a greater degree of efficiency.

Measures such as privatisation, deregulation and anything that leads to the promotion of competition can be regarded as **supply-side** measures. These are government attempts to bring about change to improve efficiency in an economy by taking action on the supply, rather than on the demand, side.

🔗 Candidates need to understand that in an examination question on privatisation, it would be appropriate to include references to both deregulation and contracting out in their answers.

Provision of information

Market failure can be caused by inadequate information. A government could therefore aim to increase the availability of appropriate information to consumers in order to try to influence their economic behaviour.

It is assumed that consumers will always aim to maximise their utility or satisfaction, but they will only be able to achieve this objective if they are in possession of the necessary information. If this information is not available, it is unlikely that they will be able to make rational decisions.

Information failure is a major cause of market failure, so a government will need to take measures to improve the accuracy and availability of information that consumers need. This will

help them make rational decisions and ensure that scarce resources are allocated as efficiently as possible. A particular example of such an approach in relation to the consumption of demerit goods, such as cigarettes, is nudge theory.

KEY SKILL

Application: a government could aim to make people as well informed as possible about the potential advantages of the consumption of merit goods (e.g. education and healthcare). It could also aim to make people as well informed as possible about the potential disadvantages of demerit goods (e.g. alcohol and tobacco). The effectiveness of such provision of information depends on the amount of money available to be spent on it and the accessibility of the communication methods used to get the information across to people.

Behavioural insights and nudge theory

- **nudge theory:** an attempt by a government to alter the economic behaviour of people in some particular way, by encouraging or discouraging consumption of certain goods or services

This is the idea that consumers do not always act in what could be regarded as a rational manner. Behavioural insights into economic decision-making stress the importance of understanding the actual behaviour of people in an economy, in contrast to the traditional approach, which emphasises the importance of acting rationally.

Nudge theory is an example of this behavioural insight into economic behaviour. The theory is based on the idea that the behaviour of consumers can be 'nudged' in a particular way, such as in relation to discouraging the consumption of demerit goods and encouraging the consumption of merit goods.

In the case of discouraging the consumption of demerit goods, such as tobacco:

» Medical evidence is very clear about the damage that tobacco can cause and yet millions of people carry on consuming tobacco. A government could decide to 'nudge' people away from the smoking of this harmful product. This could take the form of a moderate 'nudge', such as a government health warning that states that 'smoking can damage your health'. If this moderate 'nudge' did not appear to be working as effectively as had been hoped, a government could make the wording of the warning stronger, such as by changing it to 'smoking can kill' and/or by putting graphic pictures on the cigarette packets.

In the case of encouraging the consumption of merit goods, such as education:

» Positive reinforcement of the value of education could be achieved by a government informing people of the link between education and income, stressing that better-educated people generally tend to earn a higher income than less educated people.

KEY SKILLS

Application: an example of nudge theory would be discouraging the consumption of demerit goods (e.g. alcohol or tobacco).

Problem solving: nudge theory can be analysed in relation to attempts to solve certain problems in an economy, including the excessive consumption of demerit goods, such as tobacco.

Government failure in microeconomic intervention

The definition of government failure

- **Government failure:** a situation where government intervention to correct market failure does not actually improve the level of economic efficiency; such intervention may even reduce the efficiency of the allocation of scarce resources in the economy

The causes of government failure

Government failure occurs when government intervention in an economy causes an inefficient allocation of resources and a decline in economic welfare. Government failure often arises from attempts to solve market failure, but it can lead to the creation of different problems, such as other distortions or imperfections in the market.

The causes of government failure can include the following:

» **Lack of incentives:** the profit motive is usually lacking in the public sector and public sector workers may be paid less than equivalent workers in the private sector; these factors could lead to inefficiency.
» **Poor information:** a government's policy will be effective only if it has all the required information, but politicians may not always have all the information necessary to take appropriate decisions — for example, it is not easy to place a monetary value on a negative externality, such as pollution.
» **Time lags:** there may be a time lag between when a government decides to introduce a certain policy and when it actually comes into effect, by which time the economic situation might have changed.
» **Political interference:** a government decision may be taken for short-term political gain rather than for more long-term economic reasons.
» **Moral hazard:** a government could take decisions that encourage risk taking — for example, a decision to support financial institutions to avoid any of them failing and going out of business.
» **Regulatory capture**: a government could become too friendly with those it is trying to regulate.
» **Unintended consequences:** a government policy to reduce poverty through the provision of benefits could lead to a situation of 'welfare dependency'.

The consequences of government failure

There are a number of possible consequences of government failure, including the following:

» **Taxation:** a government may decide to use a progressive income tax to bring about a more equitable distribution of income, but if the top rate of tax is very high, there may be a disincentive for the higher paid to work as much; some workers may even decide to leave the country and seek employment elsewhere.
» **Employment:** a government may be reluctant to take a decision that makes people redundant — for example, if workers are relatively unproductive — because this will lead to an increase in the level of unemployment.
» **Net welfare loss to society:** when the consequences of government intervention are contrasted with the original problem that necessitated the intervention, it may be that the overall effect is a net welfare loss to society — in other words, the consequences of the government failure have actually made the situation worse than before.
» **Distortion of price signals:** one of the advantages of a free market is that it gives out price signals that lead to an efficient allocation of resources, but government intervention in a market could distort those signals so that the outcome could be an inefficient allocation of resources — for example, a government may decide to support a failing industry through subsidies when a better decision might have been to allow it to fail.

8.2 Equity and redistribution of income and wealth

The difference between equity and equality

Equity

- **equity:** the idea of fairness or justice (e.g. in relation to the distribution of income and wealth in an economy)

One objective of government microeconomic policy is the achievement of **equity**. This can be seen in terms of government policies aimed at bringing about a fairer and more equitable distribution of income and wealth.

Equality

- **equality:** the same rights and responsibilities for all the members of a group or society

Equality refers to a situation where everyone is at the same level — for example, the idea of equal life chances or of everyone having the same income and wealth. It is where there is the same status, rights and responsibilities for all the members of a group or a society.

🔗 **The role of government and the issues of equality and equity:** a government could decide to implement various policies to bring about a greater degree of equality and equity in an economy, such as through a negative income tax.

🔗 You need to be able to distinguish between equity and equality. Equity is concerned with fairness and justice whereas equality is concerned with people being the same in different respects.

The difference between equity and efficiency

- **efficiency:** the use of scarce resources in the most economical or optimal way

» **Equity:** the concept of equity was discussed above in relation to the key feature of fairness.
» **Efficiency:** another objective of government microeconomic policy is the achievement of efficiency. This involves the achievement of both productive and allocative efficiency. The attainment of such efficiency would ensure that the scarce resources in an economy were allocated in the best possible way — that is, it is concerned with optimality in the allocation and use of resources.

🔗 **Efficiency and inefficiency:** a key concept in relation to the scarcity of resources in an economy is that they should be allocated and used in the most efficient way possible.

The distinction between absolute poverty and relative poverty

- **absolute poverty:** a condition where household income is below the level necessary to maintain basic living standards in relation to food, shelter and housing
- **relative poverty:** a condition where household income is a certain percentage below the median income of a country

It is important to distinguish between absolute poverty and relative poverty:

» **Absolute poverty** is a type of poverty that occurs when the resources required for minimum physical health are lacking, defined as limited access to food, clothing and shelter. The World

Bank defines the poverty line as US$1.90 a day (using purchasing power parity, which takes into account price levels in different countries). It refers to a particular condition that is the same in every country and that does not change over a period of time

» **Relative poverty** is a type of poverty that means low income relative to others in a country. For example, it could be stated as an income that is below 50 or 60% of the median income of people in a particular country.

The poverty trap

- **poverty trap:** the situation where a person receives benefits that increase their income but means that they are no longer entitled to receive as much as was the case before

A potential problem of means-tested benefits given by a government to reduce poverty is that, as people receive money in the form of benefits, they may no longer be entitled to as much support as was the case before. This gives rise to what has been termed the **poverty trap**.

Policies towards equity and equality

- **negative income tax:** the payment of money to those people on low incomes instead of taking part of their income from them through income tax
- **universal benefits:** benefits that are provided to everyone who is entitled to them without taking into account the income of those people
- **means-tested benefits:** benefits that are provided to those people entitled to them after taking into account their income and, therefore, their need for the benefits
- **tax credits:** a form of benefit that is paid to people on low incomes to boost their income and raise their standard of living
- **universal basic income:** a payment by a government to all those entitled to it without a means test
- **transfer payment:** a form of payment to those in society who are less well off, paid for out of the revenue received from taxation

Negative income tax

Income tax has already been referred to as a way of redistributing income because an income tax is likely to be progressive. Many economies make use of progressive taxation to achieve the macroeconomic objective of a fairer and more equitable distribution of income. A progressive income tax not only takes more from a person as their income rises, but a higher proportion of that income.

A **negative income tax** can also be used to redistribute income. This involves people on low incomes receiving money from a government instead of paying income tax to it.

🔗 Candidates sometimes confuse an income tax and a negative income tax in their examination answers. An income tax involves taking money from people, whereas a negative income tax involves giving money to people.

Universal benefits

Sometimes a government will provide **universal benefits** to people that do not take into account their income.

Means-tested benefits

A government can also decide to provide benefits to those people who are less well-off through **means-tested benefits**. An example of a means-tested benefit is **tax credits**, which are paid to those people who have children or who have a job that pays a very low wage.

Universal basic income

A **universal basic income** is a government guarantee that each person in a country receives a minimum income. The idea is that the basic income will provide enough money to cover the basic cost of living.

Transfer payments

A government can decide to bring about a more equitable distribution of income and wealth through the use of **transfer payments**. This means that revenue received from taxation is used to give financial support to people, such as in the form of pensions. These payments can be regarded as worthwhile because they involve money being received by those in society who are less well off.

Inheritance and capital taxes

A government may also decide to intervene in an economy to try to bring about a more equitable distribution of wealth, as well as income. Examples of taxes that can achieve this objective are inheritance tax and capital gains tax.

KEY SKILL

Analysis: there might be examples of government failure as a result of a government deciding to intervene in a market. You need to be able to analyse the possible causes and consequences of such failure.

8.3 Labour market forces and government intervention

The demand for labour as a derived demand

- **derived demand:** where demand for the components of a product or for workers arises from demand for the final product

Labour is not demanded for its own sake, but for what it can contribute to the production process. This is known as **derived demand**.

Factors affecting the demand for labour in a firm or an occupation

- **marginal physical product:** the amount of extra output that is produced if a firm increases its input of labour by one unit

A number of factors can affect the demand for labour in a firm or an occupation, in addition to the quantity demanded of the product produced by the labour, referred to above. These include the following:

» **The price of labour and the other factors of production:** the demand for labour will depend, to some extent, on the price of labour (i.e. the wage or salary), compared with the prices of the other factors of production.
» **The productivity of labour and the other factors of production:** the demand for labour will depend, to some extent, on the productivity of labour relative to the productivity of the other factors

of production. The demand for labour is closely linked to the **marginal physical product** of labour; this refers to the additional output produced if a firm increases the labour input by one unit.

🔗 **The margin and decision making:** the concept of the margin can be analysed in relation to marginal physical product; this is the additional output produced by a firm as a result of a decision to increase the labour input by one unit.

Causes of shifts in and movement along the demand curve for labour in a firm or an occupation

The demand curve for labour is a function of the wage paid. The higher the wage rate, the lower the demand for labour; the lower the wage rate, the higher the demand for labour. The demand curve for labour, therefore, slopes downwards from left to right. Other possible factors affecting the demand for labour are assumed to be constant. Therefore, if the wage rate changes, there will be a movement along the demand curve.
A shift of a demand curve for labour occurs when there is a change in a determinant of the demand for labour, apart from a change in the wage rate. These changes include:

» changes in the productivity of labour
» changes in the skills of labour
» changes in the prices of the products produced
» changes in the demand for the products
» changes in the prices of substitutes and complements of the products

The demand curve for labour will shift inwards at a time of recession when the demand for products falls and there will be a decline in the demand for labour at each wage rate. The demand curve for labour will shift outwards in a boom when the demand for products rises and there will be a rise in the demand for labour at each wage rate.

KEY SKILL

Diagrams: the demand curve for labour slopes downwards from left to right because the higher the wage rate, the lower the demand for labour; and the lower the wage rate, the higher the demand for labour.

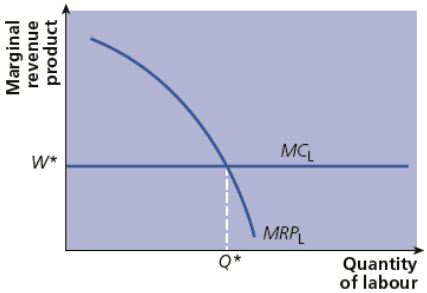
Marginal revenue product theory

- **marginal revenue product:** the extra revenue obtained by a firm as it increases its output by using an additional unit of labour

Firms are interested not only in the extra output that is produced by employing one more unit of labour, but also in the revenue obtained from selling the additional output that has been produced. The **marginal revenue product** of labour is obtained by multiplying the marginal physical product of labour by the marginal revenue received by a firm.

The derivation of an individual firm's demand for labour using marginal revenue product

F8.2 shows the profit-maximising position where the marginal cost of labour (*MC*L) equals the marginal revenue product of labour (*MRP*L) at a wage of *W** and a quantity of *Q**. In F8.2, it is assumed that a firm is in a perfectly competitive market for labour and so cannot influence the price of labour (i.e. the wage). The firm therefore regards labour supply as perfectly elastic, as shown by *MC*L.



F8.2 The labour input decision of a profit-maximising firm under a perfect competition

The margin and decision making: The concept of the margin can be analysed in relation to marginal revenue product; this is obtained by multiplying the marginal physical product of labour by the marginal revenue received by a firm deciding to employ that labour.

KEY SKILL

Diagrams: it is important to remember that the demand curve for labour is *MRP*L —the marginal revenue product of labour.

Factors affecting the supply of labour to a firm or occupation

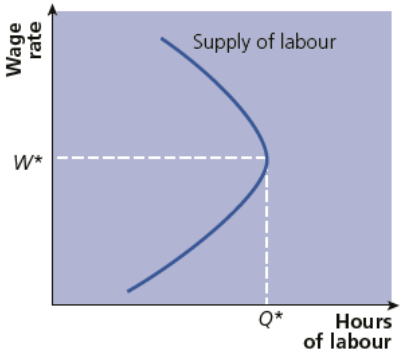
- **pecuniary advantages:** the advantages of employment that are in the form of financial rewards or benefits
- **non-pecuniary advantages:** the advantages of employment, other than the money that is gained, such as the job satisfaction gained from a particular form of employment
- **net advantages:** the overall advantages of employment, taking into account both the pecuniary and the non-pecuniary advantages

The factors affecting the supply of labour can be divided into two types:

» wage factors
» non-wage factors

For the industry as a whole, the supply curve of labour will be upward sloping from left to right because more people will make themselves available for work when the wage is increased.

For an individual worker, an increase in wages may persuade that worker to work fewer hours in order to enjoy more leisure time. This situation gives rise to a backward-bending supply curve for a particular individual. Up to a wage of *W**, the worker decides to work more hours as the wage increases, but as the wage increases above *W**, they may decide to work fewer hours.



F8.3 A backward-bending individual labour supply curve

The backward-bending supply curve of a particular worker reflects the importance of the **pecuniary advantages** of work to that employee — the reward or benefit that is paid in the form of money. Of course, there are also **non-pecuniary** reasons why people work — that is, nonwage factors. These can include:

» the working conditions
» the promotion prospects and career opportunities
» the hours of work
» the pension provision
» the availability of fringe benefits
» the facilities available at work
» the strength of vocation/job satisfaction
» the training/professional development provided

Net advantages

The balance between the pecuniary and the non-pecuniary advantages of employment is known as the **net advantages**.

Causes of shifts in and movement along the supply curve of labour to a firm or an occupation

Labour supply is defined as the number of workers willing and able to work, multiplied by the hours they are willing and able to work:

» The higher the wage rate, the more labour is supplied; the lower the wage rate, the less labour is supplied.
» The supply curve of labour, therefore, slopes upwards from left to right.
» Other possible factors affecting the supply of labour are assumed to be constant.
» Therefore, if the wage rate changes, there will be a movement along the supply curve.

A shift of a supply curve of labour occurs when there is a change in a determinant of the supply of labour, apart from a change in the wage rate. These changes include:
» the size of the working population of a country — that is, the number of people of working age who are willing and able to work — and this will be affected by such factors as changes in the retirement age and the school leaving age
» the tax and benefit levels
» net migration (i.e. the extent of immigration and emigration)
» people's preferences for work
» net advantages of work

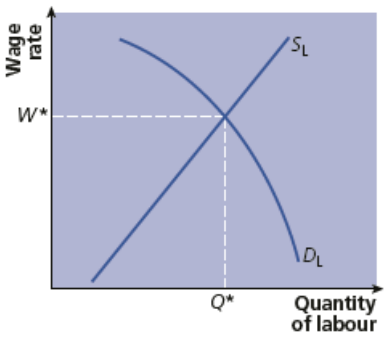
KEY SKILL

Diagrams: the supply curve of labour slopes upwards from left to right because the higher the wage rate, the greater the supply of labour; and the lower the wage rate, the lower the supply of labour.

Wage determination in perfect markets

Equilibrium wage rate and employment in a labour market

Equilibrium in a labour market is just like equilibrium in any other market — it is where demand is equal to supply. This can be seen in F8.4 where the downward-sloping demand curve for labour and the upward-sloping supply curve of labour intersect at a wage of W^* and a quantity of Q^*.



F8.4 Labour market equilibrium

🔗 **Equilibrium and disequilibrium:** situations of equilibrium and disequilibrium can be seen in labour markets, just like in any other market.

Wage determination in imperfect markets

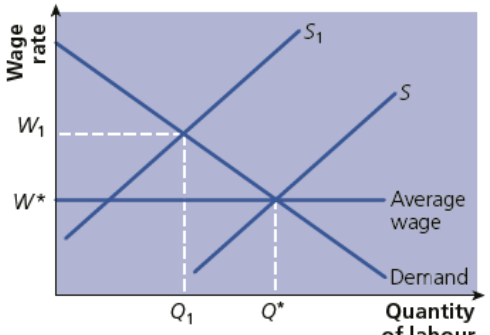
- **trade union:** an organisation of workers that is involved in collective bargaining with employers to achieve certain objectives (e.g. improvements in pay and working conditions)
- **collective bargaining:** the process of negotiation between trade union representatives of the workers and their employers on such issues as remuneration (payment) and working conditions
- **closed shop:** a requirement that all employees in a specific workplace belong to a particular trade union

A labour market may not necessarily operate as a perfect market:

» The workers may be members of a trade union, and trade unions may therefore have an influence on the process of wage determination.
» A government may decide to intervene in the process, such as through the establishment of a national minimum wage.
» There might be a single employer of labour, known as a monopsony employer.

The influence of trade unions on wage determination and employment in a labour market

It may be the case that workers belong to a **trade union** that is involved in **collective bargaining** with employers on behalf of the workers. A trade union may insist on the existence of a **closed shop** to increase its bargaining power with employers. A trade union can reduce the supply of labour — for example, by pressurising the government and/or employers into making entry into particular employment more difficult — and this will have an effect on wages. This can be seen in Figure 8.5. Without a trade union, the equilibrium position will be wage *W** and quantity *Q**.



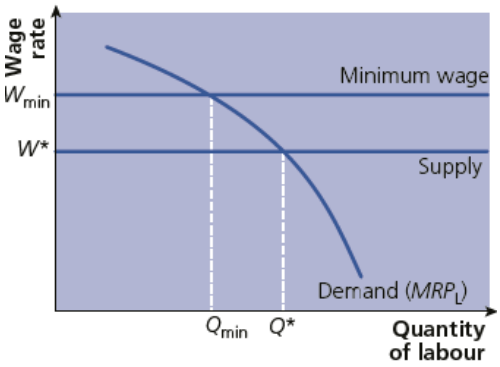
F8.5 A trade union restricts the supply of labour

However, if a trade union makes it more difficult for people to enter employment in that particular industry, supply will shift from *S* to *S*1, the wage will now b e higher, at *W*1, and the quantity of labour will now be lower, at *Q*1.

A trade union, in addition to restricting the supply of labour, could also try to increase the demand for labour, such as through the negotiation of a productivity agreement with the employer. It could also negotiate with the employer for a higher wage for the workers.

The influence of government on wage determination and employment in a labour market using a national minimum wage

» In a free market, there would be no government intervention, but governments do intervene in economies, including in labour markets.
» For example, a government might decide to introduce a minimum wage to prevent employees being paid below a certain level.



F8.6 The effect of a minimum wage on a firm in a perfectly competitive labour market

» The effect of this can be seen in Figure 8.6. Without government intervention, the equilibrium would be at a wage of *W** and a quantity of *Q**, an equilibrium determined by the intersection of demand and supply.
» A government might decide, however, that the wage of *W** is too low and so might intervene in the market by introducing a national minimum wage. This will be set above *W** at a wage of *W*min. The advantage of this is that those employed will gain a higher wage, but the disadvantage is that fewer workers will be employed (*Q*min rather than *Q**).

🔗 A common error in examination answers on this part of the syllabus is to show the minimum wage below where supply and demand intersect, rather than above this point. Candidates should understand that there would be no point in having a minimum wage below what the wage in the economy would have been without any form of intervention.

🔗 **The role of government and the issues of equality and equity:** a government could decide to intervene in a labour market in an attempt to bring about a greater degree of wage equality, such as through the establishment of a national minimum wage.

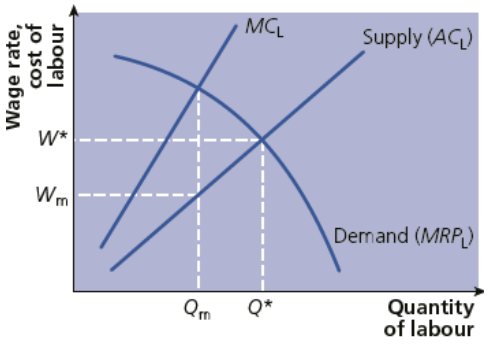
KEY SKILL

Evaluation: you will need to be able to discuss the various advantages and disadvantages of the establishment of a national minimum wage by a government in order to be able to make a judgement as to whether it is a worthwhile policy option for a government to take.

The influence of monopsony employers on wage determination and employment in a labour market

- **monopsony:** a single buyer of a product or of a factor of production such as labour

» In terms of wage determination, it has been assumed so far that the industry demand curve is made up of a number of firms in an industry. It could be the case, however, that there is just one firm in the market to employ a factor of production. Such a firm is known as a **monopsony**.
» In a competitive market, there will be many firms in an industry and so each firm must accept the prevailing market wage.
» If a firm is a monopsonist, however, the situation will be different. This can be seen in Figure 8.7. If the market were a competitive one, the equilibrium position would be where demand is equal to supply; this would be at wage *W** and quantity *Q**.
» A monopsonist, however, would be at an equilibrium position where marginal cost was equal to demand; this would lead to a lower wage of *W*m and a lower quantity of *Q*m.



F8.7 A monopsony buyer of labour

KEY SKILL

Analysis: you will need to be able to analyse the process of wage determination in imperfect markets in relation to all three possible factors — that is, the role of trade unions, the role of government and the role of monopsony employers.

The determination of wage differentials by labour market forces

If labour markets are very competitive, with identical workers and perfect mobility of labour, wages will move towards the same equilibrium level. However, wages can differ greatly and these wage differentials can be determined by a number of possible factors, including the following:

» **Variation in human capital:** this can result from differences in education and training.
» **Differences in productivity:** these can result from variations in the skills and experience of workers.
» **Discrimination:** this can exist in relation to gender, ethnicity, disability and age.
» **Compensation:** wages can compensate for risk taking, working in poor conditions or having to work unsocial hours.

These factors will influence the elasticity of supply of workers, as well as the elasticity of demand of workers. Wages will tend to be higher, the more inelastic the supply and the more inelastic the demand.

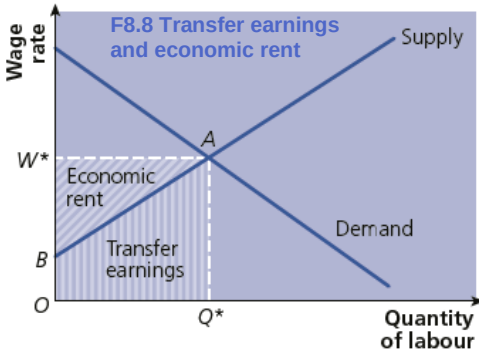
KEY SKILLS

Analysis and application: you will need to be able to analyse possible reasons for the existence of wage differentials. You will also need to be able to apply your knowledge and understanding to a particular example of wage differentials, such as between engineers and waiters.

Transfer earnings and economic rent

- **transfer earnings:** the minimum payment required to keep a factor of production in its present use
- **economic rent:** the extra payment received by a factor of production above what would be needed to keep it in its present use

Transfer earnings refer to those earnings that are the minimum that would be necessary to keep a factor of production in a particular use.



Economic rent is the additional payment that a worker receives above the transfer earnings. This can be seen in Figure 8.8, where an employee receives a wage of *W**. Some of this wage will be in the form of transfer earnings (shown by *OBAQ**) and some of it will be in the form of economic rent (shown by *BAW**).

A common error in examinations is to confuse where *economic rent* and *transfer earnings* should be shown in the diagram, as they are often labelled the wrong way round. It is important that candidates include appropriate diagrams wherever possible in their examination answers, but the diagrams need to be accurately drawn and correctly labelled.

It is possible to draw a comparison between transfer earnings and normal profit and between economic rent and supernormal profit.

KEY SKILL

Diagrams: in a demand and supply diagram where both economic rent and transfer earnings are shown, the area indicating economic rent should be above the area indicating transfer earnings.

Factors affecting transfer earnings and economic rent in an occupation

- **mobility of labour:** the degree to which labour finds it easy to move from one job to another (occupational mobility) and/or from one location to another (geographical mobility)
- **immobility of labour:** the degree of occupational immobility and geographical immobility of labour that makes a labour market less flexible than it would otherwise be
- **wage drift:** a situation where the average level of wages in an industry tends to rise faster than the supposed wage rates

The elasticity of demand and the elasticity of supply will determine the relative size of economic rent and transfer earnings in an occupation.

The amount of economic rent that a worker is able to obtain is limited by the fact that in many occupations there is not usually a high level of **mobility of labour**; in fact, in many labour markets, there is a high degree of **immobility of labour**.

There can sometimes be differences in wages in an industry in different parts of a country. This is because although there might be an officially agreed wage rate in an industry that is supposed to apply across a whole country, specific demand and supply circumstances may mean that the actual earnings of some workers in some parts of the country are higher. This situation is known as **wage drift**.

KEY SKILL

Application: a very good professional footballer will be able to gain a relatively large economic rent because both demand and supply are relatively inelastic. A cleaner, on the other hand, will receive a relatively small economic rent because both demand and supply are relatively elastic.

9 The macroeconomy

9.1 The circular flow of income

The multiplier process

The definition of the multiplier

- **multiplier:** the amount by which an increase in an injection into the circular flow of income will bring about an increase of national income in an economy

The **multiplier** is an important concept in the determination of the level of income in an economy:
» It measures the extent to which an increase in an injection into the circular flow of income brings about a magnified effect on the level of national income.
» An increase in injections, however, is also likely to have an effect on the withdrawals out of the circular flow of income in an economy.
» Each successive increase in aggregate demand will therefore become progressively less.

🔗 Even if a particular examination question does not make explicit reference to the concept of the multiplier, this does not mean that it should not be included in answers. If a question is concerned with the determination of the level of income in an economy, candidates should assume that they should refer to the multiplier in their answers.

The calculation of the multiplier

The size of an economy's multiplier will depend on how many sectors are involved.

The calculation of the size of the multiplier in an economy:

Type of economy	Calculation
Two-sector economy (households and firms)	$\frac{1}{\text{marginal propensity to save}}$ **Or** $\frac{1}{MPS}$
Three-sector economy (households, firms and government)	$\frac{1}{\text{marginal propensity to save} + \text{marginal propensity to tax}}$ **Or** $\frac{1}{MPS + MPT}$
Four-sector economy (households, firms, government and foreign	$\frac{1}{\text{marginal propensity to save} + \text{marginal propensity to tax} + \text{marginal propensity to import}}$

trade)	**Or** $\frac{1}{MPS + MPT + MPM}$

Another way of expressing the multiplier is $\frac{1}{\text{marginal propensity to withdraw}}$

KEY SKILL

Numerical skills:
the size of the multiplier in a four-sector economy is calculated by the formula of:

$$\frac{1}{\text{MPS} + \text{MPT} + \text{MPM}}$$

Therefore, if the *MPS* is 0.10, the *MPT* is 0.20 and the *MPM* is 0.30, the multiplier will be 1/0.60 = 1.667

The margin and decision making: the margin is important in relation to the multiplier because it is calculated by the formula:

$$\frac{1}{MPS + MPT + MPM}$$

The multiplier is crucial in decision making as it will have an impact on decisions taken by a government, such as in relation to expansionary or contractionary fiscal policy.

The average and marginal propensities to save, consume, tax and import

- **consumption:** the spending by consumers in an economy over a period of time
- **average propensity to consume:** the proportion of income that is spent
- **average propensity to save:** the proportion of income that is saved
- **dissaving:** a situation that can occur when consumption exceeds income, so people have to rely on savings that have been accumulated in the past
- **saving:** the amount of disposable income that is not spent on consumption
- **paradox of thrift:** the contradiction between the potential advantages and the potential disadvantages of saving in an economy
- **marginal propensity to consume:** the proportion of an increase in income that is spent
- **marginal propensity to save:** the proportion of an increase in income that is saved
- **average propensity to tax:** the proportion of income that is taxed
- **marginal propensity to tax:** the proportion of an increase in income that is taxed
- **average propensity to import:** the proportion of income that is spent on imports
- **marginal propensity to import:** the proportion of an increase in income that is spent on imported goods and services

The proportion of income that is spent on **consumption** can be measured in two ways. The first way is the **average propensity to consume**.

The average propensity to consume refers to the average proportion of income that is actually spent on buying goods and services in an economy. The proportion of income that is not spent is saved, so it is also possible to refer to the **average propensity to save**.

It is possible that consumption actually exceeds income and so would need to be financed by using past savings. This is known as **dissaving**.

Saving is generally regarded as a good thing for an individual, providing the person with the opportunity to buy something in the future. But it is also possible to view it in a negative way in terms of a society because it is a leakage or withdrawal from the circular flow of income in an economy and so could contribute to a fall in national income. This contradiction is known as the **paradox of thrift**.

The second way of measuring consumption is the **marginal propensity to consume**. Whereas the average propensity to consume is concerned with a given income, the marginal propensity to consume is concerned with a change in income and, in particular, with the proportion of that extra income that is spent. The proportion of the extra income that is not spent is saved, so it is possible to refer to the **marginal propensity to save**.

Although the main influence on consumption is the level of disposable income in an economy, there are other possible determinants. These include:

» the distribution of income and wealth
» the rate of interest
» the availability of credit
» expectations about future economic prospects

It is not only important to consider average and marginal propensities to save and consume, but also **average and marginal propensities to tax and import**. This is because knowledge of these will be required in order to calculate the multiplier in a four-sector economy.

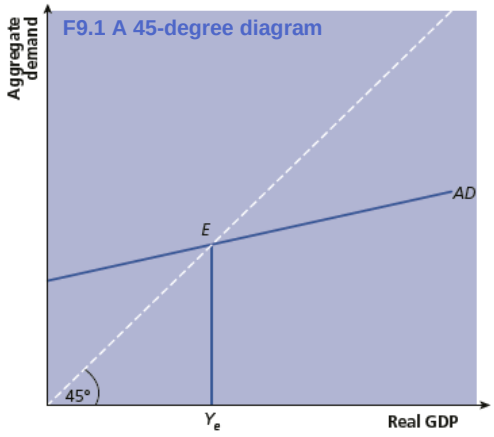
🔗 Candidates need to demonstrate they understand what is meant by the ‘paradox of thrift’. Saving is generally regarded as a good thing to do in an economy, but it is important to recognise that there may be possible disadvantages to saving, given that it constitutes a leakage or withdrawal from the circular flow of income

National income determination using the aggregate demand and income approach

There are two ways to determine the level of income in an economy. One way is by using the withdrawal (leakage) and injection approach. The other way is by using the aggregate demand and income (*AD*–income) approach.

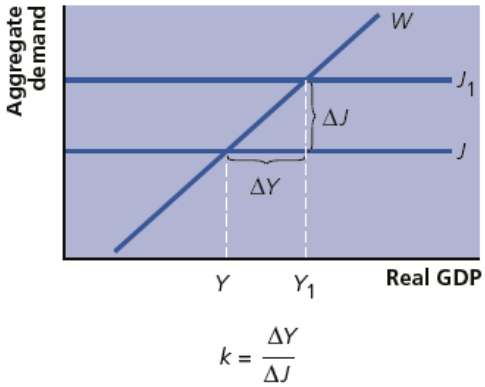
The level of income in an economy is determined at the point where aggregate demand is equal to output. Such a diagram is also known as a ‘45-degree diagram’. The economy is in equilibrium where *AD* crosses the 45-degree line, at *Y*e. Real GDP is shown on the horizontal axis and aggregate demand on the vertical axis.

F9.1 A 45-degree diagram



Calculation of the effect of changing aggregate demand on national income using the multiplier

» A change in aggregate demand can have a much greater final impact on the level of equilibrium national income than the initial change because of the multiplier effect.
» This refers to the fact that an initial injection into the circular flow of income stimulates further rounds of spending — one person's spending is another person's income.
» This will eventually lead to a bigger effect on output and employment than the initial injection.
» The operation of the multiplier can be seen in Figure 9.2. The diagram shows the effect of an increase in injections (shown by *J*) into the circular flow of income in an economy; this can be seen by the vertical distance between *J* and *J*1. The effect of this has been to increase national income from *Y* to *Y*1. The size of the multiplier (denoted by the symbol *k*) can be calculated by dividing the change in *Y* by the change in *J*.



F9.2 The effect of an increase in injections into the circular flow of income

It is possible to distinguish between a positive multiplier and a negative multiplier:

» **Positive multiplier**: when an initial increase in an injection, or a decrease in a leakage, leads to a greater final increase in real GDP.

» **Negative multiplier**: when an initial decrease in an injection or an increase in a leakage, leads to a greater final decrease in real GDP.

The components of aggregate demand and their determinants

- **autonomous savings:** savings that are not related to changes in the level of national income in an economy
- **induced savings:** savings that are related to changes in the level of national income in an economy
- **investment:** spending on capital equipment (e.g. a machine or a piece of equipment that can be used in the production process)

Aggregate demand (*AD*) refers to the total demand for, and expenditure on, all that is produced in an economy. It can be represented in the following way:

AD = C + I + G + X – M

The consumption function

The consumption function shows the relationship between consumer spending and the various factors affecting it. The main influence on consumption is the level of disposable income in an economy and the consumption function shows the relationship between income and consumption.

Autonomous and induced consumer expenditure

It is important to distinguish between autonomous consumption and induced consumption:

» **Autonomous consumption:** this refers to consumption that is not related to income — that is, consumption when income is zero.
» **Induced consumption:** this refers to consumption that is related to income — that is, as extra income is gained, some of this will be spent (the percentage of extra income that is spent is known as the 'marginal propensity to consume').

🔗 **Time:** consumption refers to household expenditure in an economy over a period of time, not just at one particular moment in time.

The savings function

The savings function shows the relationship between saving and the various factors affecting it.

Autonomous and induced savings

In the same way that it was possible to distinguish between autonomous and induced consumer expenditure, it is possible to distinguish between autonomous and induced savings:
» **Autonomous savings:** this refers to savings that are not related to income — that is, savings when income is zero.
» **Induced savings:** this refers to savings that are related to income — that is, as extra income is gained, some of this will be saved (the percentage of extra income that is saved is known as the 'marginal propensity to save').

The investment function

Investment refers to the capital expenditure by firms in an economy over a period of time, such as expenditure on factories, machinery and equipment. The investment function shows the relationship between investment spending and the various factors affecting it.
There are many influences on the investment decisions of firms, but the main determinants are:

» the rate of interest
» changes in technology
» the productivity of labour
» the cost of capital goods
» changes in consumer demand
» expectations about future economic prospects
» government policies, such as in relation to taxes and subsidies

Autonomous and induced investment; the accelerator

It is important to distinguish between two types of capital investment:

» autonomous investment
» induced investment

Autonomous investment

- **Autonomous investment:** capital investment that is not related to changes in the level of national income in an economy

Autonomous investment refers to expenditure on capital investment that is not the result of any changes in the level of national income in an economy – in other words, it is independent of any such changes. An increase or decrease in autonomous investment can be represented in a diagram by an upward shift of the aggregate demand curve for an increase in autonomous investment and a downward shift of the aggregate demand curve for a decrease in autonomous investment.

Induced investment

- **induced investment:** capital investment that is related to changes in the level of national income in an economy

Whereas autonomous investment refers to expenditure on capital equipment that is unrelated to changes in income in an economy, **induced investment** refers to expenditure on capital equipment that is directly related to changes in income. For example, a rise in national income will bring about an increase in induced investment. Induced investment is an important part of the accelerator theory of investment.

The accelerator

- **accelerator:** a way of calculating the effect of a change in national income on investment in an economy
- **capital: output ratio:** a way of measuring the amount of capital employed in the production of a given level of output

The concept of the **accelerator** theory of investment is based on the link between changes in the level of national income in an economy and changes in induced investment. It has two key features:

» It states that investment is a function of a change in national income.
» It assumes a fixed **capital : output ratio**.

It is important to understand that the accelerator is concerned with the relationship between investment and the rate of change of output — it is not the *level* of output that is important, but the *rate of change* of that output.

🔗 Candidates sometimes confuse the multiplier and the accelerator; it is important to be able to clearly distinguish between them. The multiplier shows the effect of a change in an injection, or in a withdrawal, on the level of national income in an economy. The accelerator, on the other hand, shows the effect of a change in the level of national income in an economy on the induced investment; it is a measure of how much additional capital is required to produce each extra unit of output.

Government spending

This can include government current spending on the wages and salaries of people who work in the public sector and government capital spending on investment projects, such as a new road. There are many influences on the spending decisions of a government, but the main determinants include:

» government policy commitments on particular aspects of society
» the amount of tax revenue
» demographic changes

Net exports

The concept of net exports refers to exports minus imports. The determinants of net exports can include:

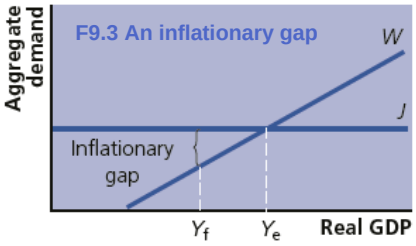
» the relative prices of a country's exports and imports
» the quality, reliability and reputation of a country's exports and imports
» exchange rate movements

The full employment level of national income and equilibrium level of national income

Inflationary gap

- **inflationary gap:** a situation where the level of aggregate demand in an economy is greater than the aggregate supply at full employment, causing a rise in the general level of prices in the economy

It is important to understand that equilibrium in an economy may not necessarily be at the full employment level of income. An **inflationary gap** shows a situation where equilibrium income is greater than the full employment equilibrium — that is, aggregate demand in an economy is greater than the full employment level of aggregate supply.

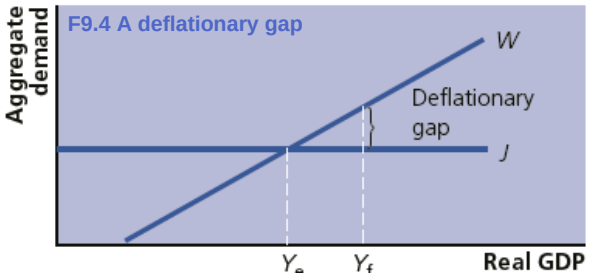


Equilibrium income is at *Ye* where the withdrawals are equal to the injections. The full employment level of income, however, is shown by *Yf*. The vertical distance between *W* and *J* at this point shows the inflationary gap in the economy.

Deflationary gap

- **deflationary gap:** a situation where the level of aggregate demand in an economy is less than the aggregate supply at full employment, causing unemployment in the economy
- **output gap:** the difference between the actual output and the potential output of an economy

Whereas an inflationary gap shows a situation where equilibrium income is greater than the full employment level of income, a **deflationary gap** shows a situation where equilibrium income is less than the full employment equilibrium — that is, aggregate demand in an economy is less than the full employment level of aggregate supply.



Equilibrium income is at *Ye* where the withdrawals are equal to the injections. The full employment level of income, however, is shown by *Yf*. The vertical distance between *W* and *J* at this point shows the deflationary gap in the economy.

It should be noted that the terms ‘inflationary gap’ and ‘deflationary gap’ can also be referred to as an **output gap**.

🔗 **Equilibrium and disequilibrium:** the concepts of equilibrium and disequilibrium are particularly important in relation to the analysis of an inflationary gap or a deflationary gap.

9.2 Economic growth and sustainability

- **economic growth:** an increase in the national output of an economy over a period of time, usually measured through changes in gross domestic product

Economic growth is defined as the increase in national output of a country over a period of time. It is usually measured in terms of a change in gross domestic product (GDP). It is important to distinguish between two types of economic growth in national output:

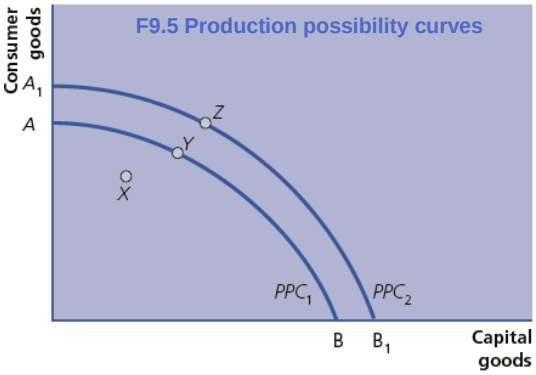
» actual growth
» potential growth

🔗 **Time**: economic growth refers to the increase in the national output of an economy over a period of time.

Actual growth versus potential growth in national output

- **actual growth in national output:** a movement from within the production possibility curve of an economy to a position on the production possibility curve, resulting from the better utilisation of the existing factors of production
- **potential growth in national output:** a shift outwards of the production possibility curve, resulting from an increase in quantity and/or quality of factors of production in an economy

» **Actual growth:** economic growth in an economy can come about by using the existing factors of production more effectively, such as by reducing the number of people unemployed. This can be shown by a movement from *within* a country's production possibility curve to a position *on* the curve. This is known as actual growth in the national output of an economy. Actual economic growth is also known as demand-side economic growth because it is affected by changes in the demand in an economy measured by an increase in real GDP over a period of time.



» **Potential growth:** it is also possible for the production possibility curve to shift outwards. This would be due to an increase in the *quantity* of the factors of production available in an economy and/or an increase in the *quality* of those factors. This is known as potential growth in the national output of an economy.

The movement from *X*, within the production possibility curve, to *Y*, on *PPC*1, shows actual growth. Potential growth is shown by the shift of the *PPC* curve from *PPC*1 to *PPC*2. The shift from point *Y* to point *Z* is the realised increase in potential growth.

🔗 It is important to make sure that you understand the distinction between actual growth and potential growth, and the distinction between a movement from inside a production possibility curve to
a position on a production possibility curve and a shift outwards of the whole curve.

Positive and negative output gaps

An output gap indicates the difference between the actual output of an economy and the maximum potential output of an economy. It is expressed as a percentage of GDP.

A country's output gap may be either positive or negative:

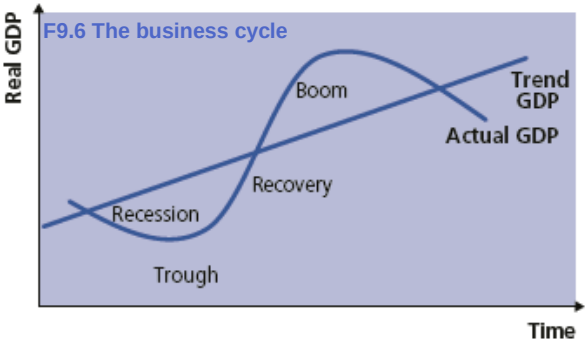
» **Positive output gap:** where an economy is outperforming expectations because its actual output is higher than the economy's recognised maximum capacity output.
» **Negative output gap:** where actual output in an economy is below the economy's full capacity output.

The business (trade) cycle

- **business (or trade) cycle:** the fluctuations in the national output of a country, involving a succession of stages or phases, including boom, recession, trough and recovery

Phases of the cycle

The **business** or **trade cycle** refers to the fluctuations in the full employment level of national output that take place in an economy over a period of time.
The cycles vary in length and seriousness, but they tend to follow the path of economic growth.



The cycle involves four stages in a sequence:

» **Boom:** a period of relatively high economic growth

» **Recession:** a period of economic downturn, defined as two successive quarters of negative GDP growth

» **Trough:** a period of low aggregate demand and relatively high unemployment

» **Recovery:** a period when the level of aggregate demand begins to increase

🔗 **Time:** the business or trade cycle refers to the fluctuations in output and employment that take place in an economy over a period of time.

The causes of the trade cycle

A number of factors can cause different phases of the cycle to occur, including:

» changes in interest rates
» changes in technology
» changes in global trade
» changes in levels of economic confidence
» changes in exchange rates
» changes in house prices
» the multiplier effect
» the accelerator effect
» changes in the level of liquidity in the financial sector
» volatility in stock market indices
» changes in fiscal policy

The role of automatic stabilisers

Automatic stabilisers are fiscal instruments that influence the rate of GDP growth and help counter swings in the business cycle:

» **Reducing the rate of economic growth:** when there is a boom and a high rate of economic growth, automatic stabilisers will help to reduce it — for example, a government will receive more tax revenue, creating a greater withdrawal from the circular flow of income.
» **Increasing the rate of economic growth:** when there is a trough and a low rate of economic growth, automatic stabilisers will help to increase it — for example, a government will increase spending on unemployment benefits, creating an injection into the circular flow of income.

The policies to promote economic growth and their effectiveness

Economic growth in an economy can be brought about by a number of factors, including:
» an increase in the number of workers
» an improvement in the quality of labour (e.g. through acquiring new skills leading to a higher level of productivity)

» a greater commitment to research and development, in terms of both invention (i.e. the discovery of new products and new methods of production) and innovation (i.e. the bringing of these inventions to the marketplace)
» an improvement in the state of technology
» investment in capital stock
» a move towards more capital-intensive production
» increased mobility and flexibility of factors of production
» a more efficient allocation of resources
» development of new markets to export to
» a reduction in taxes on company profits to allow firms to have more funds to finance investment
» an upturn in the business (trade) cycle

The effectiveness of these factors will depend on the particular economic circumstances at a given time, but key factors will be: investment, which increases productivity; education and training, which enhance the quality of human capital; technological change, which leads to the availability of better machines; and new markets for exports, which increase demand for a country's goods and services.

🔗 Don't forget that economic growth is concerned not only with the *quantity* of the factors of production used in the production process, but also with the *quality* of these economic resources.

🔗 **Progress and development:** economic growth can lead to an increase in the standard of living and quality of life of people, and so can contribute towards progress and development.

Inclusive economic growth

- **inclusive economic growth:** growth that combines increased prosperity with greater equality, creates opportunities for all and distributes the benefits of increased prosperity fairly

Impact of inclusive economic growth on equity and equality

Inclusive economic growth is a concept that advances equitable opportunities for economic participants with benefits enjoyed by every section of society. It aims to ensure that economic growth benefits everyone and therefore has a positive impact on equity and equality in an economy.

🔗 **The role of government and the issues of equality and equity:** inclusive economic growth aims to ensure that growth benefits everyone and so has a positive impact on equality and equity in an economy.

🔗 It is important that you are able to differentiate clearly between equality and equity. Although both promote fairness, equality achieves this through treating everyone the same regardless of need, whereas equity achieves it through treating people differently dependent on need.

Policies to promote inclusive economic growth

Policies to promote inclusive economic growth need to be concentrated in a number of specific areas, emphasising that it:

» takes place in the sectors in which the poor work (e.g. agriculture)
» occurs in places where the poor live (e.g. undeveloped areas with few resources)
» uses the factors of production that the poor possess (e.g. unskilled labour)
» reduces the prices of consumption items that the poor consume (e.g. food, fuel and clothing)

Policies to promote inclusive economic growth therefore need to focus on poverty reduction and eradication. They need to address inequality and enhance growth and economic inclusion by:

» expanding access to quality education
» expanding access to quality healthcare
» investing in infrastructure
» deepening financial inclusion to reach the most vulnerable
» incentivising increased female labour force participation

There are various ways to bring inclusive economic growth about, including the following:

» Progressive incomes and wealth taxes can contribute to reducing inequality without sacrificing growth.
» A universal basic income, where a government provides a guaranteed minimum income for all, has the potential to reduce poverty and inequality; it indicates the state's responsibility to support incomes in a universal way.

Sustainable economic growth

- **sustainability:** the capacity to endure; it is the potential for the long-term maintenance of well-being in an economic environment so that the interests of future, as well as present, generations are taken fully into account

- **sustainable economic growth:** a rate of growth that attempts to satisfy the needs of the present generation and sustain natural resources and the environment for future generations

Sustainability refers to the ability to use existing resources to satisfy the needs of the present generation without compromising the ability of future generations to satisfy their needs. Sustainable economic growth can therefore be defined as economic growth that takes into account the needs of future generations as well as those of the present generation.

Sustainable economic growth refers to a rate of growth that a country can maintain without creating other economic problems, especially for generations to come. It stresses the idea of a trade-off between rapid economic growth today and growth in the future. For example, rapid growth today could exhaust resources and create environmental problems for the future.

🔗 **Time:** the concept of sustainability stresses the importance of time, in that the satisfaction of the needs of the present generation should not be at the expense of those of future generations.

Using and conserving resources

The contrast between the potential benefits and costs of economic growth can be seen in relation to the use or conservation of resources. The use of resources can contribute significantly to economic growth, but it needs to be remembered that many natural resources are finite in supply — in other words, they will eventually run out.

This is why it is strongly argued in many economies that there should be conservation of resources. It is stressed that this is a more sustainable approach, taking into account the needs not only of the present generation, but also of future generations.

The impact of economic growth on the environment and climate change

Rapid economic growth today may create environmental problems for future generations, including the depletion of oil and fish stocks, and global warming.

There are clear environmental costs to economic growth:

» higher levels of resource consumption
» the depletion or loss of non-renewable resources
» greater pollution, which can cause health problems and reduce the quality of life

Some of the key facts of climate change during the last 30 years include the following:

» Global temperatures have increased by 0.5°C.
» Sea levels have risen by 10 cm.
» Carbon dioxide in the atmosphere has increased by 17%.

🔗 Sustainability is an important concept in economics, taking into account the importance of the needs of future generations and not just those of the present generation. Candidates should include a consideration of sustainability in any answer to a question on the use and conservation of resources.

Policies to mitigate the impact of economic growth on the environment and climate change

Policies to promote sustainable economic growth include the following:

» **Technology:** a government can provide financial incentives for private firms to invest in new technology (e.g. alternative sources of energy such as wind power and solar power).
» **Human capital development:** government investment in human capital by allocating more resources, and widening access to, education and training.
» **Deregulation:** a reduction of 'red tape' could encourage foreign direct investment in economies.
» **Incentivisation:** government can provide incentives to encourage people to start up their own small and medium-sized businesses.
» **Pollution permits:** these can be reduced to lower the extent of pollution in an economy over a period of time; fines can be imposed on those firms that do not respond appropriately to these permits.
» **Alternative transport to cars:** government promotion of cycling and walking, and investment in public transport, including buses, trams and trains.
» **Taxation:** a government could impose taxes on airline tickets to discourage consumption and reduce the carbon footprint, and impose carbon taxes on energy suppliers to reduce the level of carbon emissions; such initiatives would be drivers of decarbonisation.
» **Legislation:** a government could ban sales of new conventional petrol and diesel cars to encourage the purchase of greener alternatives, such as electric cars or use of bio-fuel as a source of energy, although this would require money to be spent on developing the necessary infrastructure to provide for the charging of electric vehicles.
» **Financial support:** a government could offer interest-free home and business renovation loans to drive energy efficiency; it could also offer such loans for environmentally efficient new buildings.
» **Recycling:** a government could use financial incentives to encourage recycling schemes and initiatives that will reduce demand for scarce raw materials.
» **Commitment to action on climate change:** the United Nations (UN) is encouraging countries to raise their emissions targets to include net zero emissions by 2050 (at the point of going to press, 121 out of 193 UN member countries have agreed to this target of achieving carbon neutrality) to protect people from the dangerous impact of climate change; achieving this target will require appropriate actions to be implemented.
» **Reforestation:** a government could encourage firms to commit to a policy that requires them to plant a new tree for every tree that they cut down.

Each of these policies is likely to have an impact on sustainable economic growth, reducing the potentially damaging effect of growth on the environment.

9.3 Employment/unemployment

The definition of full employment

- **full employment:** the level of employment as a result of everyone who is able and willing to work having a job

There is some debate as to what actually constitutes a situation of **full employment** in an economy, but it is generally taken to be a situation where everyone who wants a job has a job. In many economies, this would be when there was an unemployment rate of about 4%.

Equilibrium and disequilibrium unemployment

- **unemployment:** where a number of people in an economy are able and willing to work but are unable to gain employment
- **equilibrium unemployment:** a situation where jobs exist but people are unable or unwilling to take them
- **disequilibrium unemployment:** a situation where the labour market is prevented from clearing because wage rates are above their equilibrium value

Unemployment is where a number of people in an economy are able and willing to work but are unable to gain employment.

It is possible to further to distinguish between the following two broad types of unemployment:

» **Equilibrium unemployment:** this type of unemployment exists when the labour market is at equilibrium, meaning jobs exist but people are either unable or unwilling to take the jobs. Examples of equilibrium unemployment are frictional unemployment, seasonal unemployment and structural unemployment.
» **Disequilibrium unemployment:** this type of unemployment exists when the wage rate rises above equilibrium and the labour market is prevented from clearing. Examples of disequilibrium unemployment are cyclical/demand-deficient unemployment and real-wage or classical unemployment.

🔗 **Equilibrium and disequilibrium:** these concepts can be used to analyse two distinct types of unemployment, one where unemployment is due to people being unwilling to accept a job and the other where unemployment is due to wage rates being higher than their equilibrium value

Voluntary and involuntary unemployment

- **voluntary unemployment:** a situation where a worker chooses not to accept a job at the going wage rate
- **involuntary unemployment:** a situation where a worker is willing to work at the prevailing wage but cannot find a job

It is also possible to distinguish between voluntary and involuntary unemployment:

» **Voluntary unemployment:** a situation where a worker deliberately chooses not to work because of a low wage rate.
» **Involuntary unemployment:** a situation where a worker is willing to work at the market wage, but is prevented from doing so by factors beyond their control, such as a deficiency of aggregate demand or inflexibility in the labour market, especially wage rigidity. Geographical and occupational immobility of labour can be a major cause of involuntary unemployment. In an economy with involuntary unemployment, there is a surplus of labour at the current real wage.

The natural rate of unemployment

- **natural rate of unemployment:** the non-accelerating inflation rate of unemployment (NAIRU) or equilibrium unemployment; the rate of unemployment in an economy that will prevent the rate of inflation increasing

The **natural rate of unemployment** stresses the link between the level of unemployment and the rate of inflation in an economy. It is that level of unemployment which contributes towards a rate of inflation that is nonaccelerating. It is essentially an equilibrium situation where the aggregate demand for labour is equal to the aggregate supply of labour at the current wage rate; as a result of this situation of equilibrium, there is no upward pressure on the level of prices in the economy.

Determinants of the natural rate

The natural rate of unemployment is the rate of unemployment when the labour market is in equilibrium and is caused by structural, supply-side factors, such as a mismatch of skills, rather than by demand-side factors.

There are a number of determinants of the natural rate of unemployment, including the following:

» **Availability of information about jobs:** this will be an important factor in determining how quickly the people who are frictionally unemployed find jobs.
» **Level of state benefits:** relatively generous unemployment benefits may discourage people from taking jobs at the existing wage rate; however, if the level of benefits is low, this will lead to a fall in the natural rate of unemployment.
» **Skills and education:** the quality of skills and education will influence the level of occupational mobility — a better trained and better educated workforce will be more occupationally mobile and this will help to reduce the natural rate of unemployment.
» **Level of geographical mobility of labour:** the more willing people are to move to different parts of a country, the more geographically mobile they will be and this will help to reduce the natural rate of unemployment.
» **Flexibility of the labour market:** trade unions may be able to restrict the supply of labour to certain labour markets, making them less flexible, and this will increase the natural rate of unemployment.
» **Hysteresis:** a recession may cause a rise in the natural level of unemployment because when people are unemployed for a relatively long period of time, they become deskilled and demotivated, making it more difficult for them to find jobs (‘hysteresis’ refers to the delayed effects of something).

Policy implications

The way to reduce the natural rate of unemployment in an economy is to implement supply-side policies, including, but not restricted to:

» improved education and training to reduce the occupational immobility of labour
» better information about job vacancies in different parts of a country to reduce the geographical immobility of labour (mobility of labour is covered later in this section)
» greater flexibility of labour markets, such as through the reduction of the powers of trade unions

Patterns and trends in (un)employment

- **unemployment rate:** the number of unemployed people divided by the labour force

» It is important to distinguish between the level of unemployment, which is expressed as the number of people who are unemployed in a country, and the **unemployment rate**, which is expressed as a percentage.
» The unemployment rate refers to the total number of people who are unemployed in a country divided by the labour force.
» Economists are interested in discovering patterns and trends in the rate of unemployment in a country over a period of time. It is useful to establish whether the trend is upward or downward; if it is upward, the government will need to devise appropriate policies to try to reduce the rate.
» The unemployment rate of a country will fluctuate, moving up and down in line with changes in economic activity and different phases of the business (trade) cycle, but economists are interested in the long-term trend.
» Economists are interested not only in the long-term trends in employment/ unemployment, but also in the long-term changes in the patterns of employment, such as in relation to employment in the primary, secondary and tertiary sectors of economic activity.

🔗 Candidates need to distinguish between the *number of people who are unemployed* in an economy and the *rate of unemployment* in the economy. In the first case, it will be a number; in the second case, it will be a percentage.

KEY SKILL

Interpretation of data: economists are interested in discovering not only the rate of unemployment in an economy in particular months and years, but also the overall trend in that rate over a period of time.

Mobility of labour

- **geographical mobility of labour:** when a worker has the ability to move from one place to another within a country or from one country to another
- **occupational mobility of labour:** when a worker has the ability to move from one occupation to another, either in the same industry or in a different industry

Forms of labour mobility

Mobility of labour refers to the ability and willingness of labour to move from one place to another or from one occupation to another. There are thus two types of labour mobility:

» **Geographical mobility** refers to a worker's ability to move from one place to another within a country or from one country to another.
» **Occupational mobility** refers to a worker's ability to move from one occupation to another, either in the same industry or in a different industry.

Factors affecting labour mobility

There are a number of factors affecting labour mobility, including the following:

» **Education and training:** labour mobility is affected by the extent to which the labour force is educated and trained; the more a person is educated and trained, the greater their occupational mobility is likely to be.
» **Transport and communication:** a more developed transport and communication system is likely to encourage labour mobility, especially geographical mobility.
» **Job information:** the availability of appropriate information about jobs and job vacancies will impact on labour mobility, both geographical and occupational.

» **Wage differences:** differences in wages in different regions of countries, or in different countries, and in different occupations will have an influence on the extent of geographical and occupational mobility.
» **Cost of living:** the cost of living can vary a great deal between different regions of a country and between different countries, and this could have an impact on the geographical mobility of labour, especially in relation to the affordability of accommodation.
» **Immigration policy:** the ability of labour to move from one country to another may be restricted by the immigration policies of governments.

Policies to reduce unemployment and their effectiveness

A number of different policies can be used to reduce unemployment in an economy.

Fiscal policy

» Fiscal policy, used to reduce unemployment, will involve the reduction of taxation, both direct and indirect, to increase the level of consumption. Taxes on companies can also be reduced to encourage greater investment. Government expenditure can also be increased.
» A reduction in taxation and/or an increase in government expenditure will increase the level of aggregate demand in an economy and this is likely to reduce the level of unemployment.

Monetary policy

» Another possible approach to the reduction of unemployment is through the use of monetary policy.
» Interest rates could be lowered and/or the money supply increased to encourage spending in an economy.
» If the cost of borrowing is reduced, this will encourage people to spend more and save less. Also, if interest rates in an economy are lowered, this is likely to lead to a fall in the exchange rate; if this happens, it will make a country's exports more price competitive in international markets and this could lead to an increase in demand for them and therefore an increase in the derived demand for labour to produce them.

Supply-side policy

» Whereas fiscal policy and monetary policy operate to influence the level of aggregate demand in an economy, it is also possible to reduce the level of unemployment in an economy through the use of supply-side policy.
» Supply-side policies would be particularly effective in trying to reduce the natural rate of unemployment in an economy. Any policies that help markets to work more efficiently would be likely to increase the number of workers employed.
» For example, policies to make the labour market more flexible, such as restrictions on trade unions, would be likely to lead to greater employment. Education and retraining schemes would help to make workers more employable.

KEY SKILLS

Problem solving: you need to understand that there are three broad policies to reduce the problem of unemployment in an economy: fiscal policy, monetary policy and supply-side policy.

Evaluation: you will need to be able to evaluate the policies that could be used to reduce unemployment in an economy, weighing up the different benefits and limitations of the various policies, before coming to a judgement as to their relative effectiveness.

9.4 Money and banking

The definition, functions and characteristics of money

- **barter:** the direct exchange of one good or service for another
- **double coincidence of wants:** a situation in a barter system where a seller finds a buyer who wants what is being sold and where the seller also wants something that the buyer has and is willing to trade in exchange
- **money:** anything that is generally acceptable in a society as a means of payment
- **near money:** something that can perform some, but not all, of the functions of money
- **liquidity:** the extent to which a financial asset can be turned into cash (e.g. if some shares in a company are sold, the paper asset becomes money)
- **cash:** the notes and coins issued in a country that are **legal tender**
- **bank deposits:** deposits of money in accounts in financial institutions, many of which in a modern economy are in electronic form
- **cheque:** a method of payment (i.e. a means of transferring money from one account to another); it is not, however, a form of money
- **legal tender:** any form of payment that is legally recognised to settle a debt or make a payment

Before the development of money, **barter** was used. This involved the direct exchange of goods and services without the use of any form of money.

Barter, however, had a number of disadvantages:

» It needed a **double coincidence of wants** — that is, each person was required to need what the other person was offering.
» It was often very difficult to compare the value of different products.
» Some of the products would be indivisible, such as animals.
» Some products might be difficult to store during the time that a seller was looking for a buyer.

For these reasons, **money** came to replace barter. Money is defined as anything that is generally acceptable as a means of payment in an economy. The following concepts are important:

» **Near money** refers to an asset that can be immediately changed into money and can be used to settle some, but not all, debts. It can therefore perform some of the functions of money, but not all of them; it would be difficult for near money to perform the function of *medium of exchange* (see the next section).
» **Liquidity** is defined in relation to how easy it is to turn a financial asset into **cash**, with cash itself being 100% liquid. In a modern economy, some deposits are still in the form of cash, but many deposits are in the form of **bank deposits**. In this case, the money is mainly in an electronic form.
» A **cheque** can be used as a means of payment, but it is not the same as money because a cheque is not always acceptable.
» The reward for parting with liquidity is interest. If a person deposits cash in a savings account, which they can no longer use for a period of time, their reward is an additional sum of money that they receive in the future together with the amount of money originally deposited.

🔗 **Time:** if an amount of money is deposited in an account for a period of time, a reward is given in the form of interest.

🔗 Candidates should understand that the great advantage of money over barter is that it avoids the need for a ‘double coincidence of wants’. This does not mean, however, that barter has completely disappeared. In many economies, barter still exists.

The functions of money

- **medium of exchange:** the use of money as an acceptable means of payment between buyers and sellers of a product
- **unit of account (or measure of value):** the use of money to establish the value of a product
- **store of value or wealth:** the use of money to store wealth
- **standard for deferred payment:** the use of borrowed money to purchase a product now and repay the debt in the future
- **credit:** a contract agreement in which a borrower receives a sum of money now and repays the lender at a later date, usually with interest

The **four** functions of money:

Function of money	Explanation of function
A **medium of exchange**	Money works much more effectively than barter, in that money is generally acceptable as a means of payment for goods and services. This is the main reason that money is usually preferred to barter — there is no need to establish a double coincidence of wants between two people.
A **unit of account** or **measure of value**	Money enables the value of different products to be compared. This is another distinct advantage of money over barter.
A **store of value or wealth**	Wealth can be stored as money and this is much more convenient than storing items that might have been used in a barter system, such as cattle. Of course, one problem with this function of money is that inflation will reduce its purchasing power and therefore its value.
A **standard for deferred Payment**	This function of money enables people to borrow money and pay it back at a later date. This encourages **credit** and can act as an incentive to trade. Payment can be spread over a period of time — something that was much more difficult with barter.

🔗 **Time:** the function of money as a standard of deferred payment means that people can borrow money and pay it back at a later date (i.e. payment can be spread over a period of time).

The main characteristics of money in a modern economy

Acceptability	Money needs to be generally acceptable in a society if it is going to be used as a medium of exchange for buying and selling goods and services.
Portability	Money needs to be easily transported if it is going to perform its functions effectively.
Scarcity	Money needs to be relatively scarce, otherwise it will become worthless.
Recognisability	Money needs to be easily recognised; this will help to establish it and maintain people's confidence in it.
Stability of value	Money needs to be reasonably stable in value over a given period of time if people are going to have confidence in it, although inflation can negatively affect this characteristic.

Divisibility	Money must be divisible into smaller parts, or denominations, if it is going to be able to carry out its functions.
Durability	Money needs to be durable (i.e. relatively hard-wearing) over time.

🔗 Don't confuse the *functions* of money with the *characteristics* of money.

Definition of money supply

Broad and narrow money supply

- **money supply:** the amount of money available to the general public and the banking system in an economy
- **broad money supply:** a measure of the stock of money that reflects the total potential purchasing power in an economy
- **narrow money supply:** a measure of the stock of money in an economy, which is mainly cash
- **monetary base:** the cash held by the general public and by the banking system, including the balances of the financial institutions with the central bank of the country; it acts as the basis for any expansion of bank lending in an economy

The **money supply** refers to the total amount of money in an economy at any one time. It is an important macroeconomic variable. Different countries use different definitions of the money supply, but it can be broadly classified into broad money supply and narrow money supply:

» A **broad money supply** reflects the total purchasing power that is available in an economy at a particular time. It is often termed M4 and includes notes and coins plus all bank and building society deposits.
» A **narrow money supply** is mainly the cash that is available in an economy at a particular time. It is often termed M0 and includes the notes and coins held by the general public, in cash machines and in balances that the financial institutions have with a country's central bank. This narrow money supply is also sometimes known as the **monetary base**.

The quantity theory of money

- **quantity theory of money:** the hypothesis that, since $MV = PT$, and the velocity of circulation (V) and the volume of transactions (T) are constant, changes in the price level (P) in an economy are directly proportional to changes in the money supply (M)

The **quantity theory of money** shows the relationship between the money supply, the general price level and the level of output in an economy.

It hypothesises a theoretical relationship between variations in the price level and variations in the money supply. It can be expressed in terms of $MV = PT$, which is sometimes known as the Fisher equation, where:

» M is the value of the money supply
» V is the velocity of circulation
» P is the general price level
» T is the number of transactions
V and T are assumed to be constant over a period of time because it is unlikely that there is going to be a great deal of difference in the velocity of circulation or in the number of transactions in an economy from one year to the next.

In this case, *M* and *P* are directly linked. This means that if the money supply rises, people will have access to more funds, giving them greater purchasing power, and as a result of this, the general level of prices in the economy will rise — that is, a situation of inflation will exist. Persistent inflation can therefore only arise through persistent excessive growth in the money supply, which can be seen in terms of persistent outward movements of the aggregate demand curve.

The theory has been widely discussed over the years, especially in relation to whether it is correct to assume that *V* and *T* are constant over a period of time. It has also been criticised for being less of a theory and more of an identity that is necessarily true — in other words, *MV* represents total spending in an economy and *PT* represents the total money received for the goods and services. It is the same situation, but looked at in different ways.

The functions of commercial banks

Financial institutions, including commercial banks, provide an important link between borrowers and lenders, and given this position in the financial system, they are known as 'financial intermediaries'.

The functions of commercial banks include the following:

» **Providing deposit accounts:** commercial banks provide a variety of different accounts, including demand deposit or current accounts where money can be deposited and withdrawn at any time, fixed deposit accounts, where money is deposited for a fixed period of time, and various kinds of savings accounts.
» **Lending money:** commercial banks can lend money in different forms, including an overdraft, where a current account is allowed to be overdrawn up to a certain maximum amount, a loan, where a specific amount of money is lent for a particular period of time, and a mortgage, similar to a loan but usually for a longer period of time in order to buy a property.
» **Holding or providing cash, securities and equity:** commercial banks can hold or provide cash, in the form of notes and coins, and various kinds of securities, such as shares or equities in limited companies and government securities.

The reserve ratio and capital ratio

- **reserve ratio:** the proportion of the funds that a commercial bank has that it is required to maintain with the central bank and that are not available for commercial lending
- **capital ratio:** the amount of a commercial bank's capital in relation to the amount of risk it is taking

» The **reserve ratio** of a commercial bank refers to central bank regulations that establish the minimum capital reserves that a commercial bank must hold as a percentage of its deposits. The bank reserve ratio is also sometimes known as the 'cash reserve ratio (CRR)' or 'bank reserve requirement'.
» A higher proportion of reserves indicates financial soundness because a commercial bank would be better able to meet any future losses.
» The reserve ratio is a reserves to capital ratio and is calculated by reserves divided by capital.
» The **capital ratio** of a commercial bank measures the funds that it has against the riskier assets that it holds that could be vulnerable in the event of a financial crisis.
» Commercial banks are sometimes required to carry out stress tests to check that they have enough of a capital buffer to cope with any possible economic or financial circumstances.
» Commercial banks are usually required to maintain a capital ratio of at least 8%, i.e. this is a bank's core equity capital divided by its total risk-weighted assets, expressed as a percentage.

KEY SKILL

Numerical skills: the reserve ratio is calculated by reserves divided by capital. As an example, if the reserve ratio was 11% and a commercial bank had deposits of $1 billion, it would be required to have $110 million on reserve.

The objectives of commercial banks

Commercial banks have three key objectives:

» **Liquidity:** as indicated earlier in this section, this refers to the ease with which assets can be converted into cash. In relation to commercial banks, it refers to their ability to finance all of their monetary obligations to their creditors when due — these obligations can be in relation to lending, investment, the withdrawal of deposits and the maturity of liabilities.
» **Security:** commercial banks need to clearly demonstrate that they are a safe, secure and trustworthy means of storing money so that customers have confidence in them.
» **Profitability:** commercial banks are usually examples of public limited companies and, like any such company, their main aim is to make a profit for their **shareholders.**

The causes of changes in the money supply in an open economy

There are a number of different causes of changes in the money supply in an open economy.

Commercial banks as sources of credit creation and the bank credit multiplier

- **credit creation:** the process by which financial institutions are able to expand their lending by a multiple of any new deposits received; this multiple is known as the 'credit multiplier'
- **credit multiplier:** the relationship between new money created and cash deposits made

» Financial institutions are able to create 'new' money as a result of additional cash deposits. This is known as '**credit creation**'. Commercial banks know from experience that only a certain proportion of customers will want to take out money at any particular time.
» This means that a large proportion of money deposited can be loaned to new or existing customers through a process known as 'fractional reserve banking' — the requirement that such institutions only need to hold a certain percentage of their total liabilities, say 10%, in the form of cash reserves.
» The ratio of new money created to the initial money deposited is known as the '**credit multiplier**'. The value of the multiplier is calculated by 1 divided by the desired cash ratio that the commercial banks decide to hold. The smaller the cash ratio, the larger is the credit multiplier.

🔗 Despite financial institutions having a great deal of experience of how much to lend, the decision of many financial institutions to lend out too much money was a major cause of the financial crisis in 2007–08. This created a lack of confidence in financial institutions, some of which became bankrupt.

KEY SKILL

Numerical skills: the credit multiplier is calculated by 1 divided by the cash ratio. Therefore, if the cash ratio is 5%, the credit multiplier will be 1/0.05 = 20.

The role of a central bank

- **central bank:** the main bank in a country that is responsible for oversight of the banking system

A **central bank** may wish to control the ability of commercial banks to lend money, such as through open market operations. This is the process of buying or selling government securities —

that is, bonds and/or shares issued by the government. If a central bank wants to encourage bank lending, it will buy government securities.

Government deficit financing

» If government expenditure is greater than government revenue, there will be a budget deficit.
» 'Deficit financing' refers to the generation of funds to finance the deficit which results from this excess of expenditure over revenue.
» The deficit could be financed by borrowing from commercial banks or from the central bank. If the government borrows from commercial banks, for example, this will increase the banks' liquid assets, which will increase their ability to lend. This will lead to an increase in the money supply.

Quantitative easing

- **quantitative easing:** the process whereby the government, central bank or monetary authority of a country deliberately buys bonds and bills in order to increase the money supply in an economy

The process of the central bank buying existing government securities, such as bills and bonds, from financial institutions, such as commercial banks, is known as **quantitative easing**. Existing securities are bought by the central bank, leading to an increase in bank deposits, and this creates more liquidity and a greater money supply in the system. In some countries, it is known as 'credit easing'.

🔗 Candidates should understand that quantitative easing has been used by a number of countries, including the USA and the UK, to stimulate economic activity where the financial crisis of 2007–08 has led to a recession.

KEY SKILL

Evaluation: you need to be able to assess the potential benefits and limitations of quantitative easing. For example, although it can lead to an increase in aggregate demand in an economy, the greater money supply could have an inflationary impact.

Changes in the balance of payments

- **total currency flow:** an element of the balance of payments that refers to the total inflow or outflow of money which results from a country's international transactions with other countries

The **total currency flow** is a part of a country's balance of payments and shows the total inflow and outflow of money resulting from a country's international transactions with the rest of the world. If the flow is positive, it is likely to increase the foreign exchange reserves of a country. This net inflow of money into a country will have the effect of increasing its money supply.

Policies to reduce inflation and their effectiveness

There are three possible ways to reduce the rate of inflation in an economy: fiscal policy, monetary policy and supply-side policy.

Fiscal policy

A relatively high rate of inflation can be corrected by fiscal policy. For example, if a government decides to reduce expenditure and/or increase taxation, this will have the effect of reducing the

level of aggregate demand in an economy. This is a particularly useful approach to adopt when the inflation has been caused by demand-pull factors.

Monetary policy

A relatively high rate of inflation can also be corrected by monetary policy. For example, if a government decides to increase the rate of interest and/or reduce the money supply, this will have the effect of reducing the level of aggregate demand in an economy. This is also a useful approach to adopt when the inflation has been caused by demand-pull factors.

Supply-side policy

It is also possible to use supply-side policies to reduce the rate of inflation in an economy. If supply-side policies are used to make the labour market more competitive and more efficient, such as by reducing the power of trade unions or reducing unemployment benefits, this will have the effect of increasing aggregate supply. Another approach would be to use supply-side policies to increase the degree of competition in a market, such as by privatising state-owned firms and encouraging small and medium-sized businesses, which would also have the effect of increasing aggregate supply.

KEY SKILLS

Problem solving: you need to understand that there are three possible ways to solve the problem of inflation in an economy: fiscal policy, monetary policy and supply-side policy.

Evaluation: you will need to be able to evaluate the policies that could be used to reduce inflation in an economy, weighing up the different benefits and limitations of the various policies, before coming to a judgement as to their relative effectiveness — for example, supply-side policy could be successful in reducing inflation in an economy, but it might take longer and cost more money than either fiscal or monetary policy.

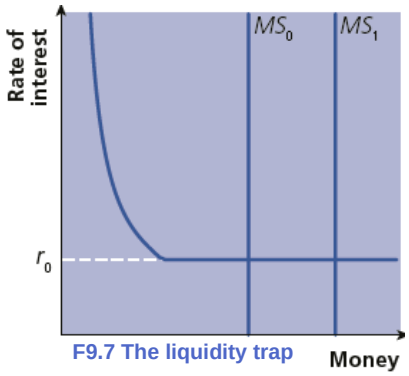
The demand for money: liquidity preference theory

- **liquidity preference theory:** the Keynesian theory of interest rate determination, based on three motives for holding money — the transactions motive, the precautionary motive and the speculative motive
- **transactions demand for money:** money that is demanded to pay for everyday purchases; it is an active balance and is interest inelastic
- **active balances:** money that is flowing through the economy, underpinning the transactions and precautionary motives for holding money
- **precautionary demand for money:** money that is demanded to pay for unexpected expenses; it is an active balance and is interest inelastic
- **speculative demand for money:** money that is demanded to buy government bonds; it is an idle balance and is interest elastic
- **idle balances:** money that is withdrawn from the circular flow of money in an economy, underpinning the speculative motive for holding money
- **yield:** the annual income obtained from a bond or a share as a proportion of its current market price
- **liquidity trap:** a situation at low rates of interest when changes in the money supply will have no effect on the rate of interest and where the demand for money is perfectly elastic

The demand for money in an economy and the determination of interest rates can be analysed through the **liquidity preference theory**.

This Keynesian approach to the demand for money and the determination of interest rates is based on three motives for holding money:

» the transactions demand for money
» the precautionary demand for money
» the speculative demand for money



F9.7 The liquidity trap

The **transactions demand for money** is where money is demanded to pay for everyday purchases. This is an **active balance** and is interest inelastic — in other words, this demand for money does not respond to changes in interest rates.

The **precautionary demand for money** is where money is demanded to pay for unexpected expenses. Like the transactions demand for money, it is an active balance and is interest inelastic.

The **speculative demand for money** is where money is demanded to buy government bonds. Unlike the transactions and precautionary demand for money, it is regarded as an **idle balance** and one that is interest elastic. An important influence on the demand for a bond is the **yield**, which is the annual income obtained from the bond as a proportion of its current market price. The price of government bonds and the rate of interest will move in opposite directions because, if the interest rate is high, this will reduce the desire to hold money. On the other hand, if the interest rate is low, there will be less of an incentive to switch out of money into other assets.

Another distinctive feature of the speculative demand for money, unlike the other two motives, is that the demand curve is downward sloping. In fact, at low rates of interest, the liquidity preference (or demand) curve becomes horizontal, indicating that a change in the money supply will have no effect on the rate of interest. At this point, the demand for money is perfectly elastic. This is known as the **liquidity trap** and it occurs when an increase in the money supply does not affect the interest rate and so does not affect investment or aggregate demand.

A shift of the money supply curve to the right would normally lead to a lowering of the rate of interest. However, when there is a liquidity trap, an increase in the money supply does not affect the interest rate, which beyond *MS*0 remains at *r*0.

Interest rate determination: loanable funds theory and Keynesian theory

Keynesian theory

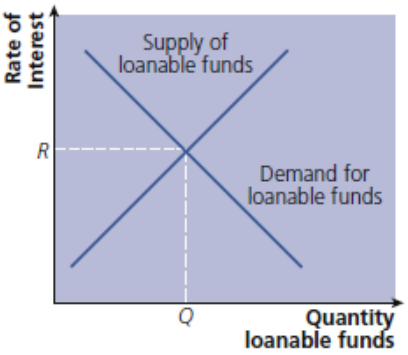
» The Keynesian theory stresses that interest is a reward for parting with liquidity for a specified period of time. According to Keynes, interest is a purely monetary phenomenon and the theory of interest is a monetary theory of interest.
» The Keynesian theory stresses that the rate of interest is determined by the demand for, and the supply of, money.
» As has already been pointed out, according to Keynesian theory, the demand for money comes about as a result of three motives: the transactions, precautionary and speculative motives.
» The supply of money is fixed and controlled by the monetary authority and is perfectly interest-inelastic.

Loanable funds theory

- **loanable funds theory:** the idea that interest rates are determined by the demand for, and the supply of, loanable funds in financial markets

» An alternative approach to the determination of interest rates is the **loanable funds theory**.

» This states that the rate of interest is determined by the demand for, and the supply of, loanable funds in financial markets.
» In other words, the rate of interest is a price and, just like any other price in an economy, it is determined by the interaction of the demand for, and the supply of, loanable funds (the supply of funds from savings and the demand for funds for investment).



F9.8 The loanable funds theory

The demand for loanable funds comes from:

» firms wanting to invest
» households wanting to buy consumer products
» a government aiming to fund a budget deficit

The demand curve for loanable funds slopes down from left to right.

The supply of loanable funds comes from savings.

The supply curve for loanable funds slopes up from left to right. the rate of interest is determined by the demand for, and the supply of, loanable funds. A rate of interest of *R* is established for a quantity *Q* of funds.

10 Government macroeconomic intervention

10.1 Government macroeconomic policy objectives

Government macroeconomic policy objectives

three macroeconomic policy objectives were stated: price stability, low unemployment, economic growth and stability of the current account of the balance of payments. It is now possible to complete the full list by adding three more areas: development, sustainability and the redistribution of income and wealth.

It is clear that governments generally have a wide range of objectives in relation to macroeconomic policy, but in most cases, they tend to focus on a combination of the following seven objectives:

» a relatively low and stable rate of inflation
» equilibrium in the balance of payments over a period of time
» a low rate of unemployment/full employment
» an appropriate rate of economic growth
» an appropriate rate of development
» growth and development that is regarded as sustainable
» the redistribution of income and wealth

🔗 **Time:** at any one moment in time, the government of a particular country may have certain priorities in relation to the achievement of these objectives, but over a period of time all governments will aim to achieve a degree of success in each of them.

10.2 Links between macroeconomic problems and their interrelatedness

It should be stressed that macroeconomic problems are not separate and distinct, but are often interrelated in different ways.

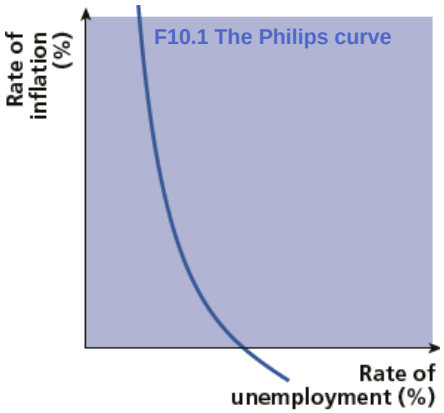
The interrelatedness of macroeconomic problems:

Relationship	Challenge
The relationship between the internal and the external value of money	It is important to distinguish between the internal and the external value of money, but it should also be recognised that the two values are interrelated. For example, if the internal value of a currency falls as a result of a high rate of inflation in a country, exports too will become more expensive. If the demand for these falls, so will the demand for the currency to pay for them. In this situation, there will be a depreciation in the external value of the country's currency.
The relationship between the balance of payments and inflation	A relatively high rate of inflation in a country will make exports more expensive, and therefore demand for them is likely to fall, but the demand for imports may remain unchanged, as many countries have a relatively high marginal propensity to consume imported products. In such a situation, the balance of payments is likely to deteriorate.
The relationship between economic growth and inflation	If fiscal and monetary policies are used in an economy to stimulate economic growth, this could lead to an increase in the rate of inflation. However, if supply-side policies are used instead, leading to a rightward shift in the long-run aggregate supply curve, this could enable economic growth to be achieved without necessarily leading to an increase in the rate of inflation.
The relationship between economic growth and the balance of payments	An increase in economic growth could result in an increase in higher real incomes and this could lead to an increase in the imports of goods and services. If this happened, it could have a negative effect on the balance of payments.

The relationship between inflation and unemployment

The traditional Phillips curve

» The relationship between the level of inflation and the rate of unemployment in an economy can be seen in the Phillips curve
» Such a curve shows the trade-off between inflation and unemployment — as the rate of inflation increases, this will be accompanied by a fall in the rate of unemployment, and as the rate of inflation decreases, this will be accompanied by a rise in the rate of unemployment.
» For example, if a government deliberately aims to bring down the rate of aggregate demand in an economy, such as through an increase in taxes or interest rates, this is likely to put some people out of work.
» It needs to be pointed out, however, that this traditional Phillips curve has now been largely discredited.



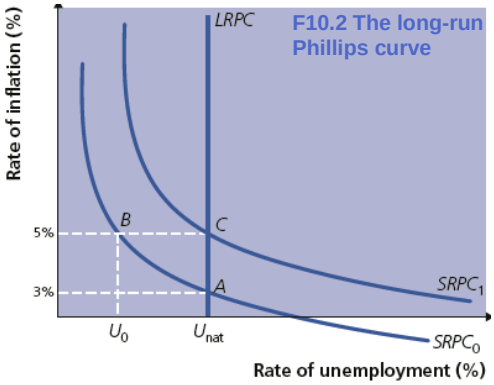
F10.1 The Philips curve

KEY SKILL

Diagrams: the Phillips curve shows the relationship between the rate of inflation and the rate of unemployment in an economy; there is a trade-off between inflation and unemployment and this is why the curve is downward sloping from left to right.

The expectations-augmented Phillips curve

» Figure 10.1 shows the Phillips curve as a downward-sloping curve.



» However, some economists have argued that there is not a trade-off between inflation and unemployment in the long run. It is argued that if actual inflation rises, expected inflation will also increase and the Phillips curve will move upwards so as to give the same expected real wage increase at each employment level. It is therefore necessary to build expectations of inflation into the relationship.
» The long-run Phillips curve therefore shows that there is a rate of unemployment that occurs when inflation is stable.
» This is known as NAIRU — the non-accelerating inflation rate of unemployment.
» Figure 10.2 shows how expectations about inflation can influence the position of the Phillips curve. *SRPC*0 shows the initial downward-sloping Phillips curve in the short run with the economy in equilibrium at point *A*. If the rate of inflation rises from 3% to 5%, there will be a movement along *SRPC*0 to point *B*.
» However, once there is a higher rate of inflation, people adjust their expectations about inflation so that the Phillips curve moves to another short run Phillips curve, *SRPC*1. Unemployment returns to *U*nat.
» The new equilibrium is now at point *C* with the same level of unemployment as before but with a higher rate of inflation.
» The long-run expectations-augmented Phillips curve is shown as a vertical straight line (*LRPC*).

🔗 **Equilibrium and disequilibrium:** the position of equilibrium and disequilibrium is seen in relation to the expectations-augmented long-run Phillips curve.

KEY SKILL

Diagrams: whereas the short-run Phillips curve is shown as a downward-sloping curve from left to right, the long-run Phillips curve can be shown as a vertical straight line.

10.3 Effectiveness of policy options to meet all macroeconomic objectives

The effectiveness of different policies in relation to different macroeconomic objectives

Fiscal policy

Fiscal policy is the deliberate adjustment of government spending and/or taxation to achieve particular macroeconomic objectives by changing the level and composition of aggregate demand.

There are two types of fiscal policy:

» **Discretionary fiscal policy:** this refers to policies that are implemented by specific one-off policy changes.
» **Automatic fiscal stabilisers:** these refer to a situation in which automatic changes in government spending and/or taxation work to help reduce the volatility of the economic cycle.

The effectiveness of fiscal policy is largely dependent on the balance between taxation and spending. Fiscal policy can be either expansionary (when there is a budget deficit) or contractionary (when there is a budget surplus).

Expansionary fiscal policy can help to reduce the level of unemployment in an economy and contractionary fiscal policy can help to reduce the level of inflation.

Taxation can fund government projects, which could help to stimulate economic growth. Progressive taxation can be especially relevant to the redistribution of income and wealth in an economy.

The advantages and disadvantages of fiscal policy

the main advantages and disadvantages of fiscal policy in relation to meeting macroeconomic objectives:

The advantages of fiscal policy	The disadvantages of fiscal policy
• Public spending can have a significant impact on the level of aggregate demand in an economy. • Public spending can help to reduce the level of unemployment in an economy. • Direct taxes (e.g. income tax) can help to redistribute income in an economy. • Indirect taxes can be targeted at altering certain kinds of behaviour (e.g. taxes on a demerit good such as tobacco). • If the spending is on infrastructure, this can help increase economic growth.	• There may be a significant time lag before a particular revenue or expenditure decision begins to take effect (e.g. between a reduction in income tax and an increase in household spending). • The potential benefit of a fiscal decision may not be fully seen because of the existence of information failure. • There may be side-effects of any fiscal policy (e.g. unemployment may be reduced, but at the cost of a higher rate of inflation). • Changing tax rates, allowances and bands is more complex than changing interest rates as part of monetary policy. • Higher taxes may have a disincentive effect on work and enterprise (e.g. higher taxes may encourage people to work less so as to have more leisure time). • A government needs to be able to estimate reasonably accurately the likely effects of any change in taxation and/or public expenditure, otherwise the desired objectives may not be achieved. • Some decisions that are justified by economic reasoning may not be taken by a government for political reasons, especially if an election is due in the not-too-distant future.

🔗 **Scarcity and choice:** a government, in carrying out fiscal policy, will need to make a choice between different areas of spending — for example, more money spent on infrastructure may mean less money spent on defence and national security, giving rise to an opportunity cost.

🔗 **Progress and development:** fiscal policy can be used to encourage economic growth and this can contribute to progress and development in an economy.

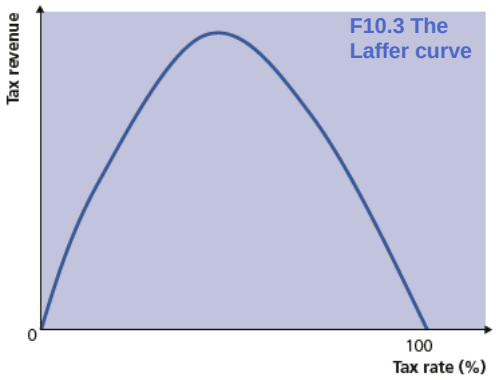
KEY SKILL

Evaluation: when analysing fiscal policy in relation to the achievement of different Macroeconomic objectives, it is necessary to evaluate the policy in terms of its various benefits and limitations. For example, fiscal policy can be effective in encouraging economic growth and reducing unemployment, but there could be a trade-off in terms of higher inflation.

Laffer curve analysis

- **Laffer curve:** a way of showing the relationship between changes in the rate of tax and changes in the total revenue gained from the tax; when the tax rate is increased, the total tax revenue first begins to rise and then falls

The effectiveness of fiscal policy can be considered in terms of Laffer curve analysis. A **Laffer curve** shows the relationship between the rate of tax and the revenue gained from it.



The Laffer curve in Figure 10.3 shows that as the rate of a tax is increased — that is, as the tax rate moves from 0 to 100% along the horizontal axis — the revenue gained from it increases up to a point. However, if the tax rate is increased further, the total tax revenue begins to fall. At these high rates of taxation, therefore, the tax is not worthwhile as it actually brings in less revenue than at the lower rates of tax.

Monetary policy

» Monetary policy is the deliberate adjustment of the money supply or interest rates to achieve particular macroeconomic objectives by changing the level and composition of aggregate demand.
» Monetary policy can be either expansionary (when there is an increase in the money supply and/or a decrease in the interest rate) or contractionary (when there is a decrease in the money supply and/or an increase in the interest rate budget).
» Expansionary monetary policy can help to reduce the level of unemployment in an economy and contractionary monetary policy can help to reduce the level of inflation.

The advantages and disadvantages of monetary policy

The advantages and disadvantages of monetary policy in relation to meeting macroeconomic objectives:

The advantages of monetary policy	The disadvantages of monetary policy
• Changes in interest rates can have a significant effect on spending in an economy, indicating that demand is interest elastic. • A central bank in many countries where it is independent from government can make decisions about changes in interest rates and/or	• There is a time lag between a change in an interest rate and the impact of that change on an economy; it is difficult to be precise about how long that time lag might be, but a number of economists have suggested that it could be as long as 18 months. • People may react differently to a change in interest rates — in other words, not everyone will have the same interest elasticity of demand. For example, relatively high-income / high-net-worth individuals in an economy are likely to have a relatively inelastic interest elasticity of demand, whereas relatively poor people are likely to have a relatively elastic interest elasticity of demand. This makes it difficult to estimate the likely effects of a change in

changes in the money supply without political interference. • Interest rates can be adjusted on a monthly basis; this contrasts with discretionary fiscal policy, which cannot be so easily changed at such regular intervals. • Interest rate changes can have a relatively immediate effect on levels of confidence in an economy.	interest rates in an economy. • Any increase in the money supply of an economy, such as through the process of quantitative easing at a time of recession, is likely to have an inflationary effect in that economy, resulting from the increase in the level of aggregate demand. • The accuracy of inflation forecasts may be relatively poor; monetary policy tries to reduce inflationary pressures (or increase them, depending on the economic conditions) before they occur, but if inflation is higher than predicted, the interest rate may be too low to control the rate of inflation effectively in an economy. • An increase in interest rates to control inflation could have a negative effect on other areas of an economy (e.g. it could have a negative impact on investment spending and therefore on economic growth, and changes in interest rates could also affect the exchange rate and the balance of payments). • The money supply may be difficult to control in practice.

KEY SKILL

Evaluation: when monetary policy is being analysed in relation to the achievement of different macroeconomic objectives, it is necessary to evaluate the policy in terms of its various benefits and limitations. For example, a reduction in the interest rate is likely to lead to an increase in aggregate demand, leading to a fall in unemployment, but there could be a time lag before this has a significant effect in an economy.

Supply-side policy

Supply-side policies can be divided into:

» **market-based policies** (e.g. reducing the size of government, lower taxes and opening up more flexible markets)

» **interventionist policies** (e.g. regional policies and education and training initiatives)

The main advantages and disadvantages of supply-side policy in relation to meeting macroeconomic objectives:

The advantages of supply-side policy	The disadvantages of supply-side policy
• Policies such as better education and training can increase aggregate supply in an economy and improve an economy's productive capacity and potential output, shifting the *LRAS* curve to the right. • Supply-side policies are of particular importance in reducing the natural rate of unemployment in an economy, especially in relation to structural and frictional	• There is no guarantee that a firm in the private sector will be more efficient than one in the public sector; privatisation may involve the need to make some people unemployed, leading to an increase in the rate of unemployment. • A privatised firm may be in a monopoly position in an economy, so all that has been achieved is a move from a state-owned to a

unemployment. • Supply-side policies can improve the level of competition, encouraging greater efficiency in markets. • Supply-side policies enable sustained economic growth to be achieved without causing a rise in inflation (depending on how the policy impacts on aggregate demand in the short run); they help reduce inflationary pressure in an economy in the long run because of the achievement of efficiency and productivity gains in the product and labour markets. • Supply-side policies can not only help to achieve a lower rate of unemployment and a higher rate of sustainable economic growth through their positive effect on labour productivity and competitiveness, but also help to improve a country's balance of payments. • Supply-side policy is, on the whole, less likely to create conflicts between the macroeconomic objectives of stable prices, sustainable economic growth, full employment and balance of payments equilibrium compared with the use of fiscal and/or monetary policy.	privately owned monopoly. • A lowering of unemployment benefits and/or a lowering of taxes might persuade more people to look for work, but this will have no effect if there are no jobs available for them. • The effects of supply-side policies can take a long time to show, as in the case of improvements to the quality of human capital through initiatives in training and education; the potential benefits of deregulation may also take a long time to have an effect in markets. • Supply-side policy can be very costly to implement, such as improvements in the provision of education and training. • Some supply-side policies may be strongly resisted if they have the effect of reducing the power of certain interest groups, such as trade unions where their influence has been reduced by certain labour market reforms. • There could be a conflict with the aim of equity, such as when supply-side policies have a negative effect on the distribution of income in an economy, at least in the short run. For example, reductions in the power of trade unions and greater privatisation could contribute to a widening of the gap between high-income/high-net-worth and low-income/low-net-worth individuals.

🔗 **Efficiency and inefficiency:** supply-side policies can lead to the more efficient use of factors of production. For example, policies such as better education and training can lead to an increase in labour productivity, increasing aggregate supply in an economy and improving an economy's productive capacity and potential output, shifting the *LRAS* curve to the right.

KEY SKILL

Evaluation: when supply-side policy is being analysed in relation to the achievement of different macroeconomic objectives, it is necessary to evaluate the policy in terms of its various benefits and limitations. For example, policies that make labour markets more flexible are likely to lower unemployment, but deregulation and privatisation could lead to a higher rate of inflation.

Exchange rate policy

» The exchange rate of an economy affects the level of aggregate demand through its impact on export prices and import prices.
» Deliberate changes in exchange rates to affect macroeconomic objectives can be regarded as a type of monetary policy.

» Changes in exchange rates have an impact on an economy through their effect on prices, in terms of both exports and imports.

The advantages and disadvantages of exchange rate policy in relation to meeting macroeconomic objectives:

Advantages of exchange rate policy	Disadvantages of exchange rate policy
• A depreciation or a devaluation can have an expansionary impact (i.e. raising the level of aggregate demand in an economy if that is what is required). • A depreciation or a devaluation can increase national output (GDP). • A depreciation or a devaluation can create jobs, leading to a reduction in unemployment. • A depreciation or a devaluation can lead to an improvement in the current account of a country's balance of payments, assuming that the *PED* for exports and imports is greater than 1. • An appreciation or a revaluation can have a contractionary impact (i.e. lowering the level of aggregate demand in an economy, if that is what is required). • An appreciation or a revaluation can reduce the rate of inflation.	• A depreciation or a devaluation, whilst lowering the price of exports, will increase the price of imports, and so it could contribute to an increase in the rate of inflation. • There will be a time-lag between any change in an exchange rate and its impact on an economy. • The scale of any change in an exchange rate may be extremely small and, in this case, it is unlikely to have much of an effect. • It is generally assumed that the *PED* for both exports and imports is likely to be relatively elastic, but this may not necessarily be the case. • An appreciation or revaluation of an exchange rate would increase the price of exports, so it could contribute to an increase in the rate of unemployment, depending on the relative size of the export sector in the economy. • The change in an exchange rate may take place at an unfavourable phase in the business (trade) cycle.

KEY SKILL

Evaluation: when exchange rate policy is being analysed in relation to the achievement of different macroeconomic objectives, it is necessary to evaluate the policy in terms of its various benefits and limitations. For example, a depreciation of an exchange rate will reduce the prices of exports and this could help to lower unemployment in an economy, but such a depreciation will also increase the price of imports and this could help to raise inflation.

International trade policy

The impact of international trade policy on the macroeconomic objectives of a government will depend on the degree of liberalisation in relation to world trade, especially the extent to which protective trade barriers, such as tariffs and quotas, have been reduced or removed.

the advantages and disadvantages of international trade policy in relation to meeting macroeconomic objectives:

The advantages of international trade policy when it promotes free trade	The disadvantages of international trade policy when it promotes free trade
• The promotion of free trade will secure market openings with trade partners. • This could lead to an increase in the level of exports to such partners. • This could reduce the level of unemployment in the export sector and encourage economic growth.	• The promotion of free trade could bring in more partners wanting to trade with a country. • This could lead to an increase in the level of imports from such partners. • This could increase the level of unemployment in an economy and lower economic growth.

The advantages of international trade policy when it promotes protectionism	The disadvantages of international trade policy when it promotes protectionism
• Infant/sunrise industries are protected, maintaining employment. • Declining/sunset industries are protected, maintaining employment. • Strategic industries (e.g. agriculture and defence) are protected, which is likely to be important to an economy.	• Resources are not allocated efficiently. • Consumers have a smaller range of products to choose from. • Economic growth may be lower than would otherwise be the case if there was no protectionism.

KEY SKILL

Evaluation: when international trade policy is being analysed in relation to the achievement of different macroeconomic objectives, it is necessary to evaluate the policy in terms of its various benefits and limitations. For example, a free trade policy could reduce unemployment and increase economic growth if an economy benefits from the greater trade, but free trade could also lead to an increase in imports, increasing unemployment and reducing economic growth. A protectionist policy could help to keep down unemployment, but economic growth might be lower than would otherwise be the case.

Problems and conflicts arising from the outcomes of these policies

It is not easy for a government to achieve success in all its policy objectives because there may often be a conflict between them. Examples of possible problems and conflicts include the following:

» A depreciation or devaluation of an exchange rate could be used to increase the demand for exports and decrease the demand for imports, thus reducing the size of a balance of payments deficit. However, if the price elasticity of demand for imports is relatively inelastic, as it often is in many countries, the demand for the more expensive imports will not be affected very much. The consequence of this is that it will contribute to inflation in an economy, both in terms of the import of raw materials and component parts and in relation to the import of finished goods.

» The possible trade-off between inflation and unemployment has already been pointed out (see section 10.1 of this chapter). Policies to control unemployment will usually have a positive effect on the rate of economic growth in an economy, but there may then be a conflict between the goal of economic growth and the need to protect the environment. Economists have certainly pointed out that a very high rate of economic growth may not be sustainable — that is, it may lead to environmental degradation and a depletion of resources that do not take the needs of future generations into account.
» If a government decided that it needed to stimulate the economy, this would have a positive effect on the level of unemployment, which would be likely to fall, and on the rate of economic growth, which would be likely to rise, but if the increase in aggregate demand were greater than the increase in aggregate supply, this would be likely to be inflationary. Moreover, if there were a relatively high marginal propensity to import, this increase in demand would be likely to worsen the economy's balance of payments position.

🔗 **The margin and decision making:** the concept of the margin is important in relation to the marginal propensity to import because if this was relatively high, an increase in aggregate demand in an economy could worsen the balance of payments position.

The existence of government failure in macroeconomic policies

It is also possible that a government may fail in its macroeconomic policies. For example:

» **Taxation:** a government could use taxation to try to bring about a more equitable distribution of income and wealth in an economy, but a number of people may decide to leave a country as a direct result of such a policy, especially if income tax becomes so progressive that people believe that they are paying too much of their income in tax.
» **Information failure:** there could be a time lag between the decision to introduce a policy and the time when that policy begins to take effect, but in the meantime, the economic situation may have changed. For example, a decision could have been taken to raise interest rates to control the level of inflation in an economy, but by the time this has begun to have an effect, inflation might have been replaced by unemployment as the main economic problem.

🔗 **Time:** there may be a time lag between a decision to introduce a particular policy and the time when that policy begins to take effect.

11 International economic issues

11.1 Policies to correct disequilibrium in the balance of payments

The components of the balance of payments accounts

- **balance of payments:** a record of all transactions linked with exports and imports, together with international capital movements; it consists of the current account, the capital account and the financial account

The **balance of payments** is a record of the transactions that a country has with the rest of the world. It shows the payments and receipts arising from this international trade. It consists of three accounts:

» the current account
» the financial account
» the capital account

There is also a balancing item.

The current account of the balance of payments

The current account of the balance of payments was discussed in Chapter 6, section 6.3.

The financial account of the balance of payments

- **financial account of the balance of payments:** records the capital inflows and capital outflows resulting from investment in different countries

The **financial account** is that part of the balance of payments which records the capital inflows into a country and the capital outflows out of a country resulting from investments. It comprises:

» direct investment (e.g. the building of a factory in another country)
» portfolio investment (e.g. the buying and selling of government securities)
» other investment (e.g. bank loans and government loans)
» reserve assets (e.g. government holdings of gold and foreign exchange reserves)

The capital account of the balance of payments

- **capital account of the balance of payments:** records transactions where there is a transfer of financial assets between one country and another; a financial asset could include the purchase of a physical asset such as land

The **capital account** is that part of the balance of payments which records capital movements, in terms of assets and liabilities, into and out of a country. These could include physical assets such as the purchase of land, the sales and purchase of patents, copyrights and royalties, and dealings such as debt forgiveness.

Net errors and omissions

- **net errors and omissions:** records those transactions in the balance of payments that go unrecorded

It is not always possible to identify all payments accurately, so there may well be some transactions that go unrecorded: these are shown in the balance of payments as **net errors and omissions**.

The balancing item

- **balancing item:** a positive or negative figure that accounts for any statistical errors in the balance of payments and ensures that the accounts, when added together, come to zero

The balance of payments should eventually balance — in other words, when all of the various parts of the balance of payments are included, the eventual figure should be zero. There may, however, be an imbalance resulting from statistical discrepancies between payments in and payments out, so the **balancing item** is used to make the balance of payments accounts balance.

The effect of fiscal, monetary, supply-side, protectionist and exchange rate policies on the balance of payments

It is now necessary to extend this discussion to include exchange rates and to consider the effect of all of them on the whole balance of payments — not only the current account, but the other accounts as well.

Fiscal policy

» If a country is experiencing a deficit in the financial account, this is not necessarily a problem as it will bring about an inflow of profits, interest and dividends in the future.
» The government could reduce taxation and/or increase public expenditure. This is likely to improve confidence in a country's economic prospects, encouraging investment and a reduction of the deficit on the financial account.
» If a country is experiencing a surplus in the financial account or the capital account, the government could increase taxation and/or reduce expenditure.

Monetary policy

» If a country is experiencing a deficit in the financial account or the capital account, the government could increase interest rates.
» This is likely to increase the flow of hot money coming into the country in search of higher interest rates.
» If a country is experiencing a surplus in the financial account or the capital account, the government could increase the growth of the money supply and/or decrease interest rates.

Supply-side policy

» If a country is experiencing a deficit in the financial account or the capital account, the government could use supply-side measures to improve the performance of the economy.
» For example, privatisation and deregulation will increase competition in markets and make domestic firms more efficient, improving quality and lowering costs.
» Increased government spending on training and education could also lead to an increase in productivity. This is likely to lead to a movement of firms and funds into the country.

Protectionist policies

» If a country is experiencing a deficit in the financial account or the capital account, the government could impose a tariff on imports.
» This would make the imported goods more expensive and so discourage their consumption, with consumers now more inclined to buy domestically produced substitutes. This would be likely to attract investment and capital into the country.

Exchange rate policies

If a country is experiencing a deficit in the current account, the government could lower the exchange rate to make its exports more competitive in world markets. However, if it was experiencing a deficit in the financial account or capital account, the government could raise the exchange rate to encourage capital movements into the country.

KEY SKILL

Problem solving: if a country is experiencing a deficit or a surplus in its balance of payments accounts (i.e. the current account, the financial account or the capital account), the government could use a variety of fiscal, monetary, supply-side, protectionist or exchange rate policies to try to overcome this problem.

The difference between expenditure-switching and expenditure-reducing policies

Two approaches can be used in an economy to correct an imbalance or disequilibrium in the balance of payments:

» expenditure-switching policies
» expenditure-reducing policies

Expenditure-switching policies

- **expenditure-switching policy:** a policy that attempts to bring about a change in the pattern of demand in an economy by reducing the demand for imports

Expenditure-switching policies are intended to switch demand away from some products and towards others. In particular, they aim to encourage an increase in the demand for exports and a decrease in the demand for imports. Different methods can be used to reduce the demand for imports. They can include:
» tariffs
» import duties
» quotas
» subsidies to domestic producers of goods that are imported
» exchange controls
» embargoes
» excessive administrative burdens on imported goods
» voluntary export restraints (VERs)

One problem associated with such policies is that they involve restraints on free trade and are therefore generally opposed by the World Trade Organization. This is why many economies decide to use expenditure-reducing policies rather than expenditure-switching policies. Another problem associated with them is that they often generate retaliation from aggrieved trade partners.

Expenditure-reducing policies

- **expenditure-reducing policy:** a policy that attempts to bring about a reduction in the level of aggregate demand in an economy in order to reduce the value of imports

Whereas an expenditure-switching policy is concerned with changing the patterns of demand for different products, an **expenditure-reducing policy** is concerned with producing a more general reduction in the demand for all products in an economy.

This will have two effects:

» The demand for imported goods will fall.
» The demand from within an economy for all goods will fall, so domestic producers will need to compensate for this by increasing the level of exports.

This process of expenditure-reducing can be carried out by various strategies, including the following:

» **Deflationary fiscal policy:** for example, through an increase in taxation and/or a reduction in government expenditure; the effect will be to create a downward multiplier effect in the economy, bringing down the level of aggregate demand.

» **Deflationary monetary policy:** for example, through an increase in interest rates and/or a reduction in the money supply.

Economists agree that the most efficient way to reduce or eliminate a balance of payments deficit is to improve the quality of the goods produced in an economy, and to reduce the unit cost of production/price, so that more people will buy them, both within the domestic market and abroad.

Many economies have adopted a variety of supply side-policies to improve the competitiveness of their domestic goods and services.

🔗 Balance of payments disequilibrium has been discussed in terms of a persistent deficit, but don't forget that disequilibrium can also refer to persistent surpluses in the balance of payments.

11.2 Exchange rates

Measurement of exchange rates

The distinction between nominal and real exchange rates

- **nominal exchange rate:** an exchange rate that is expressed in money terms without taking into effect the possible effects of inflation
- **real exchange rate:** an exchange rate that takes into account the effects of inflation in different countries
- **purchasing power parity:** the value of a currency in terms of what it would be able to buy in other countries

» The usual way to express the value of one currency in terms of another is through a **nominal exchange rate**. A nominal value is expressed in money terms, without taking inflation into account.
» However, this will not take into account the purchasing power of a nominal sum of one currency in relation to the purchasing power of another currency.
» It is necessary, therefore, to express an exchange rate as a **real exchange rate**.
» In this situation, exchange rates will be expressed in terms of **purchasing power parity** (PPP) — by taking into account price levels in different countries, exchange rates can be adjusted so as to give a more accurate comparison of the purchasing power of currencies. PPP is therefore a method of accounting for differences in the cost of living when comparing different economies.

KEY SKILL

Numerical skills: there are different ways of calculating PPP . One way is known as 'absolute PPP '. This refers to the equalisation of price levels across countries. It is calculated by dividing the price level in one country by the price level in another country. For example, if a product is priced at €375 in France and £250 in the UK, absolute PPP is calculated by €375/£250 = €1.50/£1.

Trade-weighted exchange rates

- **trade-weighted exchange rate:** a way of measuring changes in an exchange rate in terms of a weighted average of changes in other trade partner currencies

Another way of expressing an exchange rate is to relate it to changes not in one other currency, but in a number of other trade partner currencies. These different currencies are weighted to take into account their importance in international trade. As with other indices, a base year is selected and given a value of 100. The **trade-weighted exchange rate** is also called the 'effective exchange rate'.

🔗 Candidates need to demonstrate a clear understanding of the concept of purchasing power parity. This concept is important because it enables the purchasing power of a currency to be established by taking into account different price levels in various countries.

The determination of exchange rates under fixed and managed systems

Two other systems in which exchange rates are determined are fixed and managed systems.

Fixed exchange rate system

- **fixed exchange rate:** an exchange rate that is determined at a particular level by a government

Whereas a floating exchange rate system allows exchange rates to be determined by the forces of demand and supply in a free market, a **fixed exchange rate** system is where exchange rates are set at a particular level by a government. They are not determined by the forces of demand and supply.

A fixed exchange rate system is operated in the following ways:

» The central bank of a country continually buys and sells its domestic currency in order to maintain its value; it will need sufficient amounts of foreign reserves and/or gold to buy the currency when needed to maintain its value.
» The central bank can change the rate of interest to maintain a fixed exchange rate. If the exchange rate is about to fall, interest rates will be increased, whereas if the exchange rate is about to rise, interest rates will be lowered.

The advantages and disadvantages of a fixed exchange rate system

Advantages of a fixed exchange rate system	Disadvantages of a fixed exchange rate system
• It provides greater certainty in relation to economic planning and forecasting; this can encourage investment and trade. • It can help a government maintain a relatively low rate of inflation. • It provides more stability and therefore helps to limit speculation.	• It limits a central bank's ability to adjust interest rates for domestic economic reasons (e.g. in relation to economic growth). • It prevents adjustments to the exchange rate when a currency becomes overvalued or undervalued (although the fixed rate could be adjusted from time to time). • If a currency becomes extremely overvalued or undervalued, there will be a need for large devaluations or revaluations periodically, which can be more disruptive to an economy than the regular adjustment of a floating exchange rate system. • It requires a large amount of foreign reserves to support a currency when it is under pressure (i.e. when economic conditions necessitate intervention by a government, central bank or monetary authority).

Managed exchange rate system

- **managed (or dirty) float:** an exchange rate that is partly determined by the forces of demand and supply, but is also managed by a government, so that it is only allowed to float between certain parameters

Sometimes a government may allow the exchange rate to be determined by market forces to some extent, but will intervene to restrict the degree of floating between a minimum and maximum value. This is often referred to as a '**dirty float**'. A managed exchange rate system, therefore, combines elements of a floating and a fixed exchange rate system, maintaining a value which has upper and lower limits.

🔗 Candidates need to demonstrate an understanding that a managed exchange rate system is a mixture of a floating and a fixed exchange rate system.

The distinction between the revaluation and devaluation of a fixed exchange rate system

- **revaluation:** the rise in value of a currency that is fixed
- **devaluation:** the fall in value of a currency that is fixed

It is important to distinguish between a revaluation and a devaluation of a fixed exchange rate:

» When the value of a fixed exchange rate goes up, it is known as a **revaluation**.
» When the value of a fixed exchange rate goes down, it is known as a **devaluation**.

🔗 In the examination, candidates need to distinguish clearly between a revaluation/devaluation, which will occur in a fixed exchange rate system, and an appreciation/depreciation, which will occur in a floating exchange rate system.

KEY SKILL

Numerical skills: when the value of the UK£ against the US$ is reduced from £2.00 to £1.60, it is a devaluation of 20% because now a UK£ will buy less than it did before.

Changes in the exchange rate under different exchange rate systems

» These changes take place continually in response to changes in the demand for, and the supply of, a currency.
» With a fixed exchange rate system, changes in an exchange rate take place at periodic intervals when the fixed rate has become extremely overvalued or undervalued.
» If the problem is one of overvaluation, there will be a devaluation. If the problem is one of undervaluation, there will be a revaluation.
» With a managed exchange rate system, minor changes in an exchange rate will take place continually but only within certain parameters — that is, the changes will be within a narrow band.

The effects of changing exchange rates on the external economy

The Marshall–Lerner condition

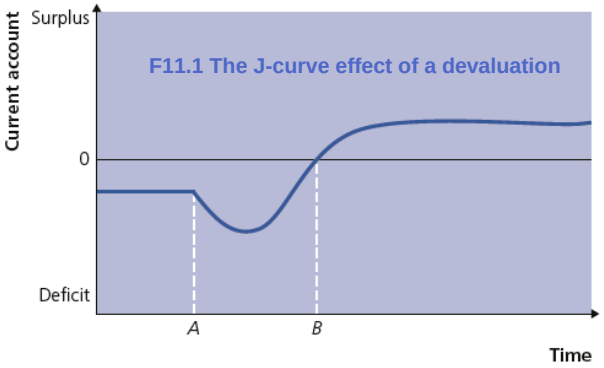
- **Marshall–Lerner condition:** the requirement that, for a depreciation or devaluation to be successful in correcting a current account deficit, the sum of the price elasticity of demand for exports and the price elasticity of demand for imports must be greater than 1

» An important factor to take into account when assessing the effect of a change in the exchange rate on an economy relates to the price elasticity of demand of imports and exports.
» A depreciation or devaluation is generally expected to bring about an increase in the demand for exports, because they are now relatively cheaper than they were, and a reduction in the demand for imports, because they are now relatively dearer than they were.
» However, the ultimate effect of these price changes will depend on the price elasticity of demand for both the exports and the imports.

» If a depreciation or devaluation is to be successful in terms of improving a current account deficit, the sum of the price elasticity of demand for exports and the price elasticity of demand for imports needs to be greater than 1. This is known as the **Marshall–Lerner condition**.

The J-curve effect

- **J-curve effect:** a situation, after a depreciation or devaluation, when the current account of the balance of payments gets worse before it gets better



F11.1 The J-curve effect of a devaluation

» A country's exchange rate can be depreciated/devalued in value in order to encourage an increase in exports and a decrease in imports.
» These effects, however, may not happen immediately and there may be a period of time when the current account deficit becomes greater before it gets better. This is known as the **J-curve effect**. It depends on the price elasticity of demand for both exports and imports.
» The reason for this J-curve effect is that buyers take time to adjust to price changes. It initially gets worse; a lower price for exports and a higher price for imports may eventually have an effect on demand, but the expected changes will not happen immediately.
» Eventually the lowering of an exchange rate may have a positive effect on the *external* economy, if the purchase of exports is encouraged and the purchase of imports is discouraged.

Time is measured on the horizontal axis and current account on the vertical axis. The current account is initially in deficit. A devaluation at time *A* will initially push the current account further into deficit because of the price inelasticity of domestic supply. The deficit will eventually reduce, but only at time *B*, when domestic firms have had time to expand their output to meet the demand for exports, will the current account move into a surplus.

🔗 **Time:** it is important to understand that the success of a depreciation in bringing about an increase in exports and a decrease in imports can be expected to take a period of time, giving rise to what has been termed a 'J-curve effect'.

🔗 Candidates need to demonstrate in their examination answers that they understand what is meant by a J-curve effect. This is the idea that there is likely to be a period of time, after a depreciation or devaluation, when the current account of the balance of payments gets worse before it gets better. It is impossible to be precise about how long this period of time is likely to be because it will vary from one country to another. You also need to make it clear that you understand that the J-curve effect depends on the price elasticity of demand for both exports and imports.

11.3 Economic development

The classification of economies in terms of their level of development

- **BRICS/BRICs:** refers to the countries of Brazil, Russia, India and China; South Africa is now usually included

Economies can be compared in terms of their level of development. It is possible to use a number of indicators to compare the characteristics of developed, developing and emerging (**BRICS**) economies.

Indicators to compare developed, developing and emerging (BRICS) economies:

Indicator	Explanation
Economic	Differences in the balance of the economic structure (i.e. the proportion of output produced in the primary, secondary and tertiary sectors of production)
Monetary	Differences in the level of income, usually measured by GNI per capita — countries can be divided into high income, middle income and low income; also, differences in percentage of GNI spent on education and health
Non-monetary	Differences in social factors (e.g. the rate of adult literacy, the number of doctors per thousand of population, the number of hospital beds per thousand of population and the quality of water)
Demographic	Differences in terms of population growth rates, birth rates, death rates, fertility rates, average age and life expectancy

🔗 It is important that you demonstrate in the examination an understanding of the fact that features of different economies tend to be rather general and that there may be exceptions. For example, there are some high income/high net worth individuals in economically developing countries and some low income/low net worth individuals in economically developed countries.

The characteristics of economically developing economies

- **birth rate:** the average number of live births per 1,000 of population of a country in a given time period, usually a year
- **death rate:** the number of people per 1,000 of population in a country who die in a given time period, usually a year
- **natural increase:** a natural increase in the population of a country is determined by the crude birth rate minus the crude death rate of a population (i.e. the difference between the number of live births and the number of deaths in a country during a year)
- **optimum population:** the number of people in a country that will produce the highest per capita economic return, given the full utilisation of the resources available
- **migration:** the movement of people from one area to another, either within a country or between countries

Population growth	There will typically be a relatively high rate of population growth, resulting from the difference between the **birth rate** and the **death rate** (i.e. the **natural increase** in a country's population); this means that, in many developing countries, the size of the population is well above the **optimum population**.
Population structure	There is likely to be a relatively high proportion of young people, creating a high dependency ratio (i.e. a high proportion of people who are reliant on others in an economy).
Income level and distribution	Income levels tend to be lower than in economically developed countries, and there is usually a great deal of inequality in the distribution of that income.
Economic structure	In many economically developing countries, a relatively large proportion of workers is employed in the primary or extractive sector and a relatively low proportion of workers is employed in the tertiary or services sector.
Employment composition	Many economically developing countries depend on the exportation of primary products and these often experience greater price variations than

	manufactured goods, creating greater instability.
External trade	Many economically developing countries depend on the exportation of primary products and these often experience greater price variations than manufactured goods, creating greater instability.
Urbanisation	The proportion of people who live in rural areas tends to be higher in economically developing countries than in economically developed countries, and where there has been rural–urban **migration**, this can put pressure on resources in the urban areas.
Dependency	Many developing countries have become dependent on economically developed countries, and the role of multinational corporations has been significant in this; although they bring employment to the economically developing countries, much of the profit made is repatriated to the home country.
External debt	Many developing countries have a high level of external debt and debt repayment is a major burden; in some countries, the debt is over 100% of the gross national product.
Social	Many economically developing countries have greater social problems than is the case in economically developed countries, as can be seen through such indicators as life expectancy, literacy rate, the number of doctors per thousand of the population and access to good-quality water.

The classification of countries in terms of their level of national income

It is also possible to classify economies in terms of their level of national income. The World Bank classifies countries according to their gross national income (GNI) per capita. It uses four categories of countries.

The World Bank classification of economies in terms of the level of national income:

Category	GNI per capita in US$
Low-income countries	Below US$1,026
Lower-middle-income countries	US$1,026–3,995
Upper-middle-income countries	US$3,996–12,375
High-income countries	Above US$12,375

The indicators of living standards and economic development

Monetary indicators

Monetary indicators include real per capita national income statistics. it was already pointed out that there are three different measurements of national income:

» gross domestic product (GDP)
» gross national product (GNP)
» gross national income (GNI)

Purchasing power parity

One issue that needs to be considered when comparing living standards and economic development in different countries is the value of the statistics in terms of what a given sum of money can buy in different countries. *

Issues of comparison using monetary indicators

- **distribution of income:** the degree to which income in a country is evenly distributed

The main method used to compare living standards and economic development in different countries is changes in real GNI. There are, however, a number of potential problems in using real GNI data.

Problems in using real GNI to compare living standards:

Problem	Explanation
The hidden, informal or underground economy	These terms refer to economic activity that is not declared and so will not be included in the official GNI data (e.g. it may be illegal).
Non-marketed goods and services	The GNI data are likely to be reasonably accurate when the vast majority of economic output is recorded through market transactions, but in many economies there is a lot of economic activity that is not marketed (i.e. there is no price attached) and so it goes unrecorded.
The difficulties of measuring government spending	In some countries, it can be difficult to measure spending by the government because it is not always easy to value the output of something that is not sold in a market (e.g. defence).
Sustainability	A country may have a relatively high rate of economic growth, but this will not necessarily be regarded as good if the needs of future generations are not sufficiently taken into account (i.e. if the growth is not sustainable because natural resources may be depleted and an unacceptable level of pollution created).
Unequal **distribution of income** and wealth	Real GNI can be divided by a country's population to give an indication of the average standard of living, but this average may be misleading if there is an unequal distribution of the GNI.
The contrast between increases in consumer goods and increases in capital goods	A rise in living standards will come about as a result of an increase in consumer goods available, but in the short run there may be an increase in the production of capital goods which will eventually help to make this possible.
The quantity and quality of output	GNI statistics measure the quantity of a country's output, but they do not take into account the quality of that output.
The effect of exchange rate changes	When comparing standards of living between people in different countries, it is necessary to take into account the effect of changes in the exchange rates between countries which would otherwise distort any comparison; economists achieve this through the use of purchasing power parities.

🔗 **Progress and development:** sustainability needs to be taken into account when evaluating the progress and development of societies.

Non-monetary indicators

Although monetary indicators are used to compare living standards and economic development in different countries, non-monetary indicators can also be used. However, there are also problems with using these.

Problems in using non-monetary indicators to compare living standards:

Problem	Explanation
The level of literacy	The literacy rate can vary a great deal between countries and those countries with a relatively low rate of literacy may not produce very accurate data.
Working hours and working conditions	The GNI data record the quantity of output produced in different countries, but they do not take into account the way in which that output is produced — a country's output may have increased substantially, but this may have been at a cost of a significant increase in working hours and a deterioration in working conditions.
Political freedom	Although not directly an economic consideration, it could be argued that political freedoms and civil/human rights need to be taken into account when assessing the quality of life in different parts of the world.

Composite indicators

- **Measure of Economic Welfare:** a broader measure of economic welfare than real GNI per capita; it takes into account such aspects as the value of childcare and looking after the sick and elderly, any depletion of natural resources and changes in the natural environment
- **Human Development Index:** a composite measure that combines life expectancy, average income in the form of GNI per capita (PPP US$), and years of schooling
- **Multidimensional Poverty Index:** this replaced the Human Poverty Index in 2010 and is made up of three dimensions and ten indicators; like its predecessor, it focuses on the extent of deprivation in different countries

differences in real per capita GNI are the most established way of comparing standards of living in different countries, but they are not the only method used to make such comparisons. Other approaches have been used, many of which use both monetary and non-monetary elements. They are called composite indicators because they use more than one indicator and compile these different indicators into a single index. Composite indicators include the following:

» **Measure of Economic Welfare (MEW):** this approach includes such elements as leisure hours, crime rates and levels of pollution.
» **Human Development Index (HDI):** this takes into account gross national income per head (at PPP in US$), life expectancy and years of schooling.
» **Multidimensional Poverty Index (MPI):** like the Human Poverty Index, which it has replaced, this focuses on the extent of deprivation in different countries.

The Human Development Index

The HDI uses three dimensions:

Dimension	Indicators
Education	Mean years of schooling Expected years of schooling
Life expectancy	A life expectancy component is calculated
Standard of living	Gross national income (GNI) per capita adjusted for purchasing power parity

Each country is classified into one of four groups according to the HDI figure:

» Low: HDI score between 0.00 and 0.54.
» Medium: HDI score between 0.55 and 0.69.
» High: HDI score between 0.70 and 0.79.
» Very high: HDI score above 0.80.

🔗 Candidates should note that the Human Development Index was revised in 2010 to use GNI per capita rather than GDP per capita to measure income.

KEY SKILL

Numerical skills: the HDI is constructed from three separate indices: one-third from the Education Index, one-third from the Education Index and one-third from GNI.

The Measure of Economic Welfare

The MEW incorporates four dimensions:

» the value of GDP
» the value of leisure time
» the value of unpaid work
» the value of environmental damage

The Multidimensional Poverty Index

The MPI uses ten indicators in three dimensions:

Dimension	Indicators
Health	• Child mortality • Nutrition
Education	• Years of schooling • Child school attendance
Living standards	• Electricity • Sanitation • Drinking water • Type of floor • Type of cooking fuel • Ownership of assets

🔗 Candidates need to be clear about what is, and what is not, included in these various composite measures. For example, the quality of drinking water is *not* included in the Human Development Index, but *is* included in the Multidimensional Poverty Index.

In examination questions on the comparison of living standards and economic development in different countries, it would be helpful if candidates brought as many of the various indicators of the standard of living between countries as are relevant into their answers.

The quality of life

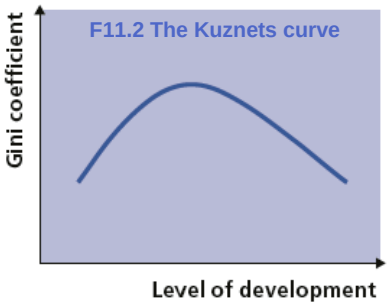
- **quality of life:** a wider concept than the standard of living used to compare living conditions in different countries; it could take into account such criteria as the number of doctors per thousand of population, the quality of drinking water and the average size of school classes

Economists have tended to focus on the concept of standard of living, especially when GDP per capita was the main measure. There has, however, been greater use of the concept of **quality of life** in recent years, reflecting the use of these composite measures. The concept of quality of life is a broader concept than standard of living and takes into account a wider range of criteria.

The Kuznets curve

- **Kuznets curve:** a curve that shows that as an economy develops, economic inequality first increases and then decreases

The **Kuznets curve** shows that as an economy develops over time, economic inequality first increases and then decreases. This idea was first put forward by the economist, Simon Kuznets, in the 1950s. It has since been heavily criticised for being overly simplistic.



the Kuznets curve with the Gini coefficient on the vertical axis and the level of development on the horizontal axis. The curve shows that, to begin with, inequality increases, but over a period of time, as a country develops, the degree of inequality begins to decrease.

🔗 **Time:** the Kuznets curve shows that as an economy develops over time, economic inequality first increases and then decreases.

The comparison of economic growth rates and living standards

It is possible to make comparisons of economic growth rates and living standards:

» over time
» between countries

Over time

» Real GNI data can be analysed over time to recognise trends in economic growth in a country. These may relate to significant upturns or downturns at particular times which will have important effects on standards of living.
» For example, many countries experienced a downturn after the financial crisis of 2007–09. Economic growth rates can be averaged out so that comparisons over time can be made.
» Changes in productivity rates can also help to explain economic growth rates over time. Another factor to take into account when comparing living standards over time is whether a country is experiencing a persistent trade deficit and, if it is, by how much.

» If a country is experiencing a persistent trade deficit, it means that it is living beyond its means. It is also necessary to consider changes in the level of public debt and especially whether the level of debt is rising or falling as a percentage of GNI.

Between countries

There may be similarities in economic growth rates between countries, especially in the long run, for a number of reasons, including the following:

» **Technological developments:** a key factor in determining similarities in economic growth rates and living standards is the development and implementation of new technologies. All countries have been able to benefit to varying degrees from similar improvements in technology.
» **The role of multinational companies:** multinational companies operate in many economies all over the world and this can contribute to improvements in productivity in all countries where they are located.
» **Global shocks:** all countries, to varying degrees, are subject to the same global shocks, such as when there is a significant rise in oil prices or an international financial crisis.

However, there may be differences in economic growth rate trends between countries, especially in the short run, for a number of reasons, including the following:

» **Government demand management policies:** governments may pursue different macroeconomic policies — for example, one country may pursue deflationary policies of higher taxation and lower spending, while another country pursues reflationary policies of lower taxation and higher spending.
» **Industrial relations:** some countries may experience better industrial relations than other countries, contributing to higher productivity growth.
» **Entrepreneurial culture:** some countries may experience a greater level of dynamism and innovation than other countries, perhaps because of differences in the level of government support for the development of an entrepreneurial culture.

🔗 **Time:** it is important to be able to make comparisons of economic growth rates and living standards in a particular country over a period of time.

11.4 Characteristics of countries at different levels of development

Population growth and structure

It is important to have an understanding of the measurement and causes of changes in:

» the birth rate
» the infant mortality rate
» the death rate
» net migration

Birth rate

The birth rate is the number of individuals born into a population in a given length of time. It is measured by the annual number of live births per 1,000 inhabitants. Possible causes of changes in the birth rate are as follows:

» The need for large families, so that children can work at an early age, can contribute to an increase in the birth rate.
» Improvements in education and healthcare can contribute to a decrease in the birth rate.

Death rate

The death rate (also known as the ‘mortality rate’) is the number of deaths in a particular population during a particular period of time. It is measured by the annual number of deaths per 1,000 inhabitants.

Possible causes of changes in the death rate are as follows:

» An epidemic or pandemic can contribute to an increase in the death rate.
» Better food and nutrition and improved health services can contribute to a decrease in the death rate.

Infant mortality rate

- **infant mortality rate:** the number of children dying under one year of age divided by the number of live births that year

The **infant mortality rate** (also known as the ‘infant death rate’) is the number of deaths in a group younger than 1 year of age. It is measured as the annual number of deaths of those under 1 year old per 1,000 live births.

Possible causes of changes in the infant mortality rate are as follows:

» An increase in infectious diseases can contribute to an increase in the infant mortality rate.
» Better healthcare can contribute to a decrease in the infant mortality rate.

Net migration

- **net migration:** the number of immigrants into an area minus the number of emigrants from an area

Net migration is the difference between the immigrants coming into, and the emigrants leaving, a country. It is measured by the number of immigrants minus the number of emigrants over a given period of time, usually 1 year.

Possible causes of changes in net migration are as follows:

» An increase in the number of immigrants coming into a country, with the number of emigrants constant, will increase net migration.
» A decrease in the number of immigrants coming into a country, with the number of emigrants constant, will decrease net migration.

The optimum population

The optimum population has a number of features:

» It refers to the ideal number of population that a country should have, after taking into account its available resources.
» It is not a fixed or rigid population size, but is variable depending on possible changes in the quantity and quality of resources and in the level of technology.
» It is therefore defined as that size of population enabling maximum per capita output to be achieved, accompanied by the highest possible standard of living within a given set of economic and technological conditions.

The level of urbanisation

- **urbanisation:** the process of the movement of people from rural areas to urban areas

» A key aspect of population structure is the level of **urbanisation** in a country. It needs to be pointed out that the definition of what constitutes an urban area can vary between countries.
» Urbanisation describes the movement of people from rural areas, i.e. hamlets and villages, to urban areas, i.e. towns and cities.
» As countries become more developed, the level of urbanisation increases. Today, 55% of the world's population live in urban areas, a proportion that is expected to increase to 70% by 2050.

Income distribution

Calculation of the Gini coefficient and Lorenz curve analysis

The Gini coefficient

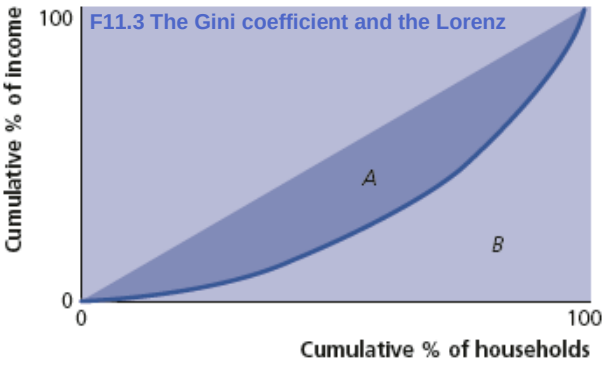
It is necessary to consider how income distribution can be measured in an economy.

» The Gini coefficient is a way of measuring the extent of inequality in the distribution of income in an economy. It can also be used to measure wealth distribution in an economy.
» The Gini coefficient was discussed in Chapter 3, section 3.3. It is measured as the ratio of the area between the diagonal line of total equality and the Lorenz curve (see below) to the total area under the diagonal. The bigger this area, the more unequal is the distribution of income.
» In this way, the Gini coefficient measures the extent to which the distribution of income in an economy differs from the position of total equality. The lower the value of the coefficient, the more even is the distribution of income. The higher the value of the coefficient, the less even is the distribution of income.

The Lorenz curve

- **Lorenz curve:** a graphical representation of the degree of inequality in the distribution of income in an economy

The **Lorenz curve** is a graphical representation that shows the extent of inequality in the distribution of income in an economy. The more unequal the distribution of income in an economy, the more divergent the Lorenz curve will be from the diagonal line of total equality.
the Gini coefficient is obtained by calculating the ratio of the area between the equality line (the straight line in the diagram) and a country's Lorenz curve (area *A* in the diagram) to the whole area under the equality line (area *A* + *B* in the diagram).



KEY SKILL

Interpretation of data: the lower the value of the Gini coefficient, the more even is the distribution of income. The higher the value of the Gini coefficient, the less even is the distribution of income.

Economic structure

Employment composition

- **industrialisation:** the process by which an economy is transformed from primarily agricultural to one based on the manufacture of goods
- **deindustrialisation:** a reduction in the size or share of the manufacturing sector in an economy

Production in an economy can take place in three sectors: the primary sector, the secondary sector and the tertiary sector. As a country becomes more developed, its employment structure and composition changes:

» **Primary sector:** mechanisation reduces the need for agriculture and raw material extraction, so the percentage of workers employed in the primary sector decreases.

» **Secondary sector:** as a country becomes more **industrialised**, many workers leave jobs in agriculture, forestry and fishing (i.e. the primary sector), and take up jobs in the growing secondary sector (e.g. manufacturing). However, as a country develops further, there will to some extent be an element of **deindustrialisation**, and the percentage of people employed in the secondary sector, having first increased, will then decrease.

» **Tertiary sector:** the decline in the proportion of people working in the primary and secondary sectors will continue and the proportion of people working in the tertiary sector will increase (e.g. service occupations such as education and finance).

Indonesia is an example of this change in employment composition as it has become more developed.

Changes in the Indonesia employment composition (%):

Sector	1920	2020
Primary sector	51	33
Secondary sector	32	22
Tertiary sector	17	45

The pattern of trade at different levels of development

As a country becomes more developed, its pattern of trade changes:

» Developing countries tend to have economies that are largely based in the primary sector, so most of their exports are agricultural and/or other primary products, and most of their imports are manufactured products.
» Developed countries tend to have economies that are largely based on manufacturing and services, so most of their exports are from the secondary and tertiary sectors.
» This pattern of trade at different levels of development tends to favour developed, rather than developing, countries in that the price of agricultural products tends to be lower than the price of manufactured products.

11.5 Relationship between countries at different levels of development

International aid

Forms of international aid

- **aid:** the process of economically developed countries providing financial support to economically developing economies

International **aid** refers to any form of assistance by one country or institution to another. It is usually given by economically developed to economically developing economies and it can come in various forms. These can include:

» bilateral (i.e. involving just two countries)

» multilateral (i.e. involving a number of countries and/or agencies, such as the World Bank)
» humanitarian assistance (e.g. aid for medical or educational reasons)
» emergency or short-term aid (e.g. after a sudden disaster)
» conditional or tied aid (e.g. a grant or loan with conditions attached)
» unconditional or untied aid (e.g. a grant or loan without any conditions)
» charitable aid (e.g. funded by donations from organisations such as Oxfam)

Reasons for giving international aid

There are a number of reasons for giving international aid, including the following:

» It helps countries fight diseases.
» It helps countries to respond to disasters and humanitarian emergencies.
» It helps countries affected by a hunger crisis.
» It helps countries to improve their health and education systems.
» It helps to save the lives of people living in poverty.
» It helps countries around the world to improve their infrastructure (e.g. in relation to water supply).

The effects of international aid

As has been indicated, international aid can assist the social and economic development of a country and/or help it to respond more effectively to a disaster.

There have, however, been a number of problems in relation to international aid, including the following:

» Some aid has been used in investments which have become relatively unsuccessful. For example, some major construction projects, such as dams, have been criticised for not being in the long-term interests of particular countries.
» There have been some examples of the aid money being spent on defence, rather than on, say, education, healthcare or improving the water supply.
» Some aid has actually been counterproductive and has made the situation worse — for example, supplies of large amounts of food to some countries have reduced the prices in the local economy, lowering incomes and making it difficult for local farmers to survive.
» Some of the aid has been ‘tied’ in some way — that is, there are conditions attached to the aid, so the receiving country has not been totally free to decide how to spend the money.
» In some cases, there has been corruption in the receiving country, with the result that much of the money has been kept in the hands of a few, rather than spread across the whole economy.
» While in some cases the interest rate charged on the aid has been at a lower rate — that is, concessional — in many cases it has been at market rates of interest.
» The loan repayments have, in many cases, been difficult for the economically developing countries to pay, although some debts have been rescheduled or cancelled; organisations such as Jubilee 2000 and Make Poverty History have been active in trying to persuade lenders to cancel, or at least reschedule, the debts.

The importance of international aid

International aid has been of importance for a number of reasons, including the following:

» It helps the economy of the receiving country to grow.
» It helps to improve the level of employment in a country.
» It helps a country to improve its balance of payments position.
» It allows for experts to come into a country to provide technical advice and assistance.
» It can save the lives of millions of people living in poverty around the world.

» It helps to address health, education, infrastructure issues.
» It helps to address humanitarian emergencies.
» It helps to promote international trade.
» It helps to redistribute global wealth.

Trade and investment

The theory of comparative advantage indicates that all countries can benefit from free trade as long as there are differences in opportunity cost ratios. International trade will therefore lead to an increase in world output and this will ultimately lead to an improvement in the quality of life and economic welfare in all countries, including economically developing economies.

Economically developing economies, however, have experienced particular problems, which means that they have not gained from international trade as much as might have been expected. These problems include the following:

» Whereas economically developed economies have tended to specialise in the production of manufactured goods, economically developing economies have tended to specialise in the production of primary products; some economically developing economies rely on a single commodity for over half of their export earnings. The problem is that the prices of primary products have, on the whole, declined relative to the prices of manufactured goods, although many economically developing economies are now much less dependent on the exports of primary products than used to be the case.
» Primary production is likely to be more volatile in terms of supply conditions, with the effect that prices for primary products are usually less stable than for manufactured goods; a key issue is that the demand for, and the supply of, primary products is more price inelastic than is the case with secondary products.
» The demand for many primary products, such as different types of food, is more income inelastic than demand for manufactured goods.
» Much of the investment in economically developing economies has come from multinational companies; much of their profit is repatriated back to their home countries.

Some economically developing economies, given these problems, have actually introduced a number of different forms of trade protection to bring about import substitution in their economies, aiming to replace imports with domestically produced goods. Some economically developing economies, however, recognising the enormous potential offered by free trade, have taken a different approach and have aimed for an export-led strategy in order to benefit from free trade.

The role of multinational companies

Definition of a multinational company

- **multinational company (MNC):** a firm that operates in different countries

A **multinational company (MNC)** is one that has facilities and other assets in at least one country other than its home country. It generally has factories and/or offices in different countries.

The activities of multinational companies

MNCs are involved in a range of activities in different countries, including:

» producing and selling goods and services
» exporting goods and services
» making significant investments in those countries
» buying and selling licences

🔗 note that it is not enough for a company just to sell products in other countries to be described as a multinational; it needs actually to have operations located in other countries, such as factories.

The consequences of multinational companies

The advantages and disadvantages of MNCs:

The advantages of MNCs	The disadvantages of MNCs
• They can bring jobs, creating more income, which can contribute to a local multiplier effect, reducing unemployment. • They can provide more choice for consumers, leading to a higher standard of living. • They can lead to an increase in tax revenue for the domestic government, such as from taxes paid on profits. • They can bring technical knowledge that could lead to higher levels of productivity. • They can stimulate economic growth. • If the output produced is exported, this could lead to an improvement in the current account of the balance of payments.	• They may use capital-intensive methods of production, which means that not many local jobs are created. • The jobs that are created may be relatively unskilled (they are sometimes known as 'screwdriver' jobs). • Much of the profit may be repatriated back to the MNC's home country and not reinvested in the local economy. • They may damage the environment. • They may try to influence the government of the country, leading to the possibility of corruption. • When a MNC leaves a country, the void created could be worse in the long run than if it had never been there.

Foreign direct investment

- **foreign direct investment (FDI):** investment by a company, with a head office in one country, in another country in the form of a factory or distribution outlet

Foreign direct investment (FDI) is investment in a country by an investor from another country. It is a form of entry into a foreign market. The Organisation of Economic Cooperation and Development (OECD) defines control when the foreign investor owns 10% or more of the business.

The consequences of foreign direct investment

The advantages and disadvantages of foreign direct investment:

Advantages of foreign direct investment	Disadvantages of foreign direct investment
• Foreign expertise can be an important factor in the improvement of the existing technical processes in a country.	• It can be a hindrance to domestic investment because the investor is investing elsewhere than in the investor's home country.

• It can help improve the quality of products and processes in particular sectors of an economy. • It can help in the creation of jobs and so reduce the level of unemployment. • It provides a source of external capital for a country that can improve its level of economic development. • It provides a source of tax revenue to a government.	• Political issues can quickly arise, making FDI potentially very risky. • A government could decide to take control of the investment for political purposes. • FDI can sometimes have an influence on exchange rates to the advantage of one country and the detriment of another. • Investment in some countries could be relatively expensive and it may be more expensive to locate production abroad than to export goods.

External debt

The nature of external debt

External debt (also known as 'foreign debt') refers to the portion of a country's debt that has been borrowed from foreign lenders, including commercial banks, governments and international financial institutions.

External debt represents the amount that a particular country owes to other countries. It includes both public sector debt and private sector debt. It also includes both short-term liabilities (e.g. loans that need to be repaid in the near future, such as within 1 year) and long-term liabilities (e.g. loans that need to be repaid over a longer period of time).

The causes of external debt

The causes of the existence of external debt include:

» outstanding loans to foreign private sector financial institutions, including the outstanding interest
» payments due to international organisations, such as the International Monetary Fund
» outstanding payments for a balance of payments deficit

The consequences of external debt

It is not easy to be precise when deciding when a country's external debt has become a problem. One of the key issues is the ability of a country to meet the interest payments on the external debt. These payments will need to be met from:

» foreign currency earnings from exports
» foreign currency reserves
» gold reserves
» further borrowing

The existence of external debt can be a significant obstacle to the economic growth and economic development of a country. The repayment of the debt can become a major burden and there is an opportunity cost involved: the funds used to repay the debt could have been used in other ways that would have been more productive, such as spending on healthcare and education.

It is as a result of the problems caused by the repayment of external debt that the following institutions have been established:

» the International Monetary Fund
» the World Bank

The role of the International Monetary Fund

- **International Monetary Fund (IMF):** set up in 1944 to promote international trade through such measures as providing financial support in the form of a loan, which will help a country to overcome, or at least reduce, a deficit in the balance of payments; it has 190 member countries

A number of international organisations have been established since 1944 to encourage free trade and to reduce the extent of trade protectionism in the world market. One of these is the **International Monetary Fund (IMF)**. It was set up to secure international monetary cooperation, to stabilise currency exchange rates, and to expand international liquidity through access to hard currencies.

The IMF aims, in particular, to:

» reduce the extent of global poverty
» encourage international trade
» secure financial stability
» promote sustainable economic growth
» promote high employment
» foster global monetary cooperation

The IMF achieves these aims by overseeing economic development, lending and capacity development. It has played a significant role in stabilising exchange rates and thereby facilitating international payments. It has also helped to enforce monetary discipline among its member countries.

The role of the World Bank

- **World Bank:** set up in 1944 mainly to provide finance to developing countries to help various kinds of projects; it has 189 member countries

Another international organisation established since 1944 is the **World Bank**. The World Bank Group comprises five different institutions, of which one key department is the International Bank for Reconstruction and Development.

The World Bank aims to:

» provide low-interest loans, interest-free credit and grants to middle-income and low-income countries to reduce the extent of poverty
» improve the health, education and infrastructure facilities of countries
» modernise the financial sector, agriculture and natural resources and environmental management of different countries

The World Bank has set two goals for the world to achieve by 2030:

» end extreme poverty by decreasing the percentage of people living on less than US$1.90 a day to no more than 3%
» promote shared prosperity by fostering the income growth of the bottom 40% for every country

11.6 Globalisation

The meaning of globalisation and its causes and consequences

The meaning of globalisation

- **globalisation:** the process whereby there is an increasing world market in goods and services, making an increase in multilateral trade more likely; it has been made possible by a number of factors, including progress in trade liberalisation

» **Globalisation** is a process of interaction and integration among people, companies and governments worldwide.
» It refers to the increase of trade around the world, especially by large companies producing and trading goods in many different countries.
» It involves an increase in international competition through the development of a global free market, with companies able to gain a competitive advantage. Developments in transport and financial deregulation, in particular, have helped to facilitate this.
» It refers to the free movement of goods, services and people across the world and involves a process that enables financial and investment markets to operate internationally.

The causes of globalisation

Globalisation describes an economic interdependence of countries around the world fostered through the development of trade liberalisation and free trade. The reduction of protectionism, involving the removal or reduction of tariffs and other import controls, has been a significant cause of globalisation. It is also the result of deregulation and improvements in communications technology.

The consequences of globalisation

The benefits of globalisation	The limitations of globalisation
• It has led to the increased spread of products, technology, information and jobs across national borders. • It has created new jobs and enhanced economic growth through the international flow of goods, capital and labour. • Companies have been able to reduce costs by manufacturing abroad and have been able to gain access to millions of new customers. • Significant improvements in standards of living have been achieved in many countries, as free markets and free trade are a means of reducing the extent of poverty in the world. • Global competition in markets has led to the production of better-quality products at lower prices.	• Job creation and economic growth have not been distributed evenly across industries or countries. • Specific industries in certain countries have suffered disruption or collapse as a result of increased competition. • It has benefited large multinational companies, but its impact remains mixed for certain workers and small businesses around the world, in both developed and developing countries. • It can have a significant detrimental effect on the environment. • It can create a domino effect where a crisis or economic downturn in one country has a severe impact on many other countries in the world.

The distinction between a free trade area, a customs union, a monetary union and full economic union

Economic integration

There are different forms of economic integration in various parts of the world and it is important that these different types of integration are clearly distinguished.
The different forms of economic integration are:

» a free trade area
» a customs union
» a monetary union
» an economic union

Free trade area

- **free trade area:** a group of countries that promote free trade between themselves, but that retain a separate set of trade barriers against other countries
- **common external tariff:** where all of the countries in an economic organisation impose the same tariff towards other countries outside the organisation

A **free trade area** refers to a situation where a number of countries come together to trade freely between each other, but where each of these countries maintains its own trade barriers with other countries outside the free trade area — in other words, there is no **common external tariff** barrier between the member countries and other countries.

Examples of free trade areas include:

» the European Free Trade Area (EFTA), comprising Iceland, Liechtenstein, Norway and Switzerland
» the North American Free Trade Area (NAFTA), comprising the USA, Canada and Mexico, although this has now been replaced by the United States–Mexico–Canada Agreement (USMCA)
» the South Asian Free Trade Area (SAFTA), comprising Afghanistan, Bangladesh, Bhutan, India, The Maldives, Nepal, Pakistan and Sri Lanka

Customs union

- **customs union:** a group of countries that promote free trade between themselves, and that impose a common external tariff on imports from countries outside the area

A **customs union** refers to a situation where a number of countries come together to trade freely between each other, and in addition to this, the countries establish a common external tariff against all other countries that are not members of the customs union.

The European Union was a customs union at one point in its development before it achieved full economic integration. The Southern Africa Customs Union (SACU), comprising Botswana, Namibia, South Africa, Lesotho and Eswatini, is an example.

Monetary union

- **monetary union:** a group of countries that decide to bring their economies closer together through the adoption of a single currency

A **monetary union** refers to a situation where a group of countries come together and adopt a single currency. They may also adopt a number of common monetary policies to support the operation of that currency.

Economic union

- **economic union:** a group of countries which agree to integrate their economies as much as possible through various rules, laws, policies and regulations

» An economic union refers to a situation where a number of economic policies, rules and regulations are established which affect all the member countries.
» An economic union may have a single currency, but it is not necessary that all member countries agree to use that currency.
» The European Union (EU) consists of 27 countries (a 28th country, the United Kingdom, withdrew from the European Union on 31 January 2020), but only 19 of them use the single currency, the euro. The total population of the European Union is 450 million.
» The Eurasian Economic Union (EEU) came into existence on 1 January 2015 and comprises five countries: Russia, Belarus, Armenia, Kyrgyzstan and Kazakhstan. It has a population of 183 million. There is no common currency and each of the five countries uses its own currency.

🔗 A number of candidates seem to believe that all countries in the European Union use the single currency, the euro. This is not the case. In 2021, 8 of the 27 member countries did not use the euro. Latvia became the 18th member country of the Eurozone in 2014 and Lithuania became the 19th member country of the Eurozone in 2015.

🔗 Candidates often confuse a free trade area with a customs union. Both types of integration encourage free trade between the member countries, but the key difference between them is that in a free trade area the member countries retain their own trade barriers with countries outside the area, whereas in a customs union the member countries adopt a common external tariff towards countries outside the union.

Trade creation and trade diversion

- **trade creation:** the creation of new trade as a result of the reduction or elimination of trade barriers
- **trade diversion:** where a certain amount of trade is lost as a result of the imposition of trade barriers

» When a trading bloc has been created, business will begin to take advantage of the opportunities arising from free trade between the member countries.
» As a result of the establishment of the trading bloc, there will be fewer trade barriers. New markets are likely to open up and businesses will try to take advantage of the opportunities offered by this situation. There will be increased specialisation and more trade. The overall effect, therefore, is likely to be one of **trade creation**.
» However, when a new trading bloc is established, there is the possibility that there will be some degree of **trade diversion**.
» The member countries of the bloc may want to shift to buying more from member countries and to buying less from non-member countries. The absence of trade barriers between the member countries of the bloc encourages them to trade with each other rather than with countries outside of the bloc.

Learn by heart!